The Dissertation Committee for An-Shih Liu
certifies that this is the approved version of the following dissertation:

Three Essays on Empirical Studies of Consumer Behavior

Committee:

Kenneth Hendricks, Supervisor

Randal Watson, Supervisor

Eugenio Miravete

Russell Cooper

Romana Khan

Three Essays on Empirical Studies of Consumer Behavior

by

An-Shih Liu , B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2007

Acknowledgments

I would like to thank my supervisors, Dr. Hendricks and Dr. Watson for their encouragement and guidance through the development of this dissertation. I would like to thank my Mom, my Dad, and my sisters, Pei-Shih and Li-Shih as well for their moral support during the time I was pursuing the doctor's degree in the United State. This dissertation would not have been completed without their help.

Three Essays on Empirical Studies of Consumer Behavior

Publication No. ______________

An-Shih Liu , Ph.D.
The University of Texas at Austin, 2007

Supervisors: Kenneth Hendricks
Randal Watson

This dissertation is an empirical study of demand and supply in differentiated products markets using supermarket scanner data on two particular product categories - canned tuna and hot-breakfast cereals. First, I study the impact of retailers' price promotions on consumer demand and retailer profits in the canned-tuna product category. Since canned tuna is storable, I examine whether consumers stock up during sales. The results suggest that only a limited amount of stockpiling exists in this product category. Since inventory is not very important, consumer demand is thus modeled by a static demand model with a random-coefficients-nested-logit specification, which is estimated by the Markov Chain Monte Carlo method. The unit-sales decomposition results show that on average 36% of the demand response to price promotions comes from brand-switching, so market expansion effects due to consumers switching from the outside good and to higher quantities usually dominate the brand-switching effect. Using the demand estimates, I compute

optimal retail prices assuming that stores are local monopolists and choose prices to maximize static category-level profits. I find that regular prices at "high-low" stores are typically at or slightly below the optimal prices, but that regular prices at "every-day-low-price" stores are substantially below the optimal prices. These results suggest that retail price levels and price promotions are more likely related to local market conditions such as retail competition.

In addition, I study the effects of store-brand (SB) entry on the demand elasticities of incumbent national brands (NB), consumers' substitution patterns for national and store brands, and the implications for consumer welfare in the hot-breakfast-cereals product category. A random-coefficients model of consumer demand is estimated by the generalized-method-of-moments approach. The empirical findings are: (1) After the entry of SB's, demand becomes more elastic for non-imitated NB's, and either more elastic or shows no change for imitated NB's; (2) in general, substitution patterns for NB's and SB's are asymmetric, i.e., when the prices of their favorite products increase, most NB buyers tend to substitute to other NB products, but SB buyers will substitute to the corresponding imitated NB's; (3) the increase in consumer surplus due to SB entry is trivial for an individual consumer, but the aggregate benefit could be quite substantial.

Table of Contents

Chapter 1

An Empirical Study of Price Promotions: The Case of Canned Tuna

1.1 Introduction

Why do retailers hold sales, and why do they hold them at particular times and frequencies? The aim of this chapter is to study these questions in the context of supermarket retailing of a particular product - canned tuna. As with many or most goods sold at supermarkets, this product is frequently observed to be on price promotion ("on sale"). In the data used below, defining a sale to be a temporary discount of at least 20% off the normal price, I find that most supermarkets had at least one brand of tuna on sale in at least half the weeks over a sample period of several years. Since the frequencies and magnitudes of sales for a particular product can have a major impact on revenues, it is of interest to study the role of these sale characteristics in the store's overall profit maximization problem.

My empirical approach is to first estimate a static model of individual consumer demand using scanner data on household purchases. Consumers' simultaneous choices of brand and quantity are modeled using a random-coefficients-nested-logit approach. The data come from the ERIM data set, which contains household purchases of grocery products in Springfield, MO, over 123 weeks from 1985 to 1987. Using these data I estimate a demand model by the Markov Chain Monte Carlo method, which is computationally much faster than likelihood methods and which permits recovery of both population-level and individual-level parameters.

I then use the demand estimates to perform two kinds of counterfactual experiments. In the first I estimate the sources of the boost in sales quantities during price-promotion periods. That is, following a stream of related research in the marketing literature, I ask how much of a sales boost (or "bump") comes from: (a) consumers switching from other product categories or no consumption (the "outside good"); (b) consumers switching from other brands in the same category; and (c) increased quantities purchased by existing consumers of the brand. This decomposition is of interest because the relative magnitudes of these components tells us something about the implied profitability of the price promotion from the respective points of view of retailers and manufacturers. If most of the sales bump comes from consumers in category (b), then the promotion is apparently benefiting a particular manufacturer more than the retailer, since the former realizes increased quantities at the expense of a rival brand, whereas the latter just observes a shift in sales from one inventory item to another (and that other item may be sold at a lower price). (In other words, behavior of type (b) represents "cannibalization" from the retailer's point of view.) To the extent that the promotion does not reduce its sales of goods in other categories, the retailer may do better when consumers of types (a) and (c) are relatively more important, since those types represent purchases of tuna that would not have been made in the absence of the promotion.

While the above decomposition exercise provides indirect evidence about the retailer's profitability of sales, in my second counterfactual exercise I investigate this profitability directly, using estimates of retailer marginal costs. The ERIM data do not contain wholesale prices, so I derive these marginal-cost estimates from the well-known Dominick's supermarket scanner database. Under plausible assumptions the average acquisition costs from Dominick's can be used as proxies for the missing costs in the ERIM data. Given these estimated wholesale prices and the results of my demand estimation I calculate optimal retail prices for each store in the data, assuming that a store is a local monopolist who chooses prices each week to maximize static profits. I then compare these predicted optimal prices with those actually observed in the data, both during and outside of sale periods.

With respect to my results on the decomposition of a sales bump, I find that the composition varies considerably depending upon the brand on promotion and upon the magnitude of price discount. In general, however, the ordering of the magnitudes of the three effects is, from highest to lowest, purchase incidence (i.e., switching from the outside good), brand switching, and purchase quantity. On average, across brands and discount depths, the increase due to brand switching is only about 36%, so market-expansion effects usually dominate the business-stealing effect. In particular, price promotions are effective in inducing purchases of canned tuna from previous non-purchasers. Using the demand estimates, I then back out wholesale prices and run a counterfactual experiment to study retailer pass-through of a manufacturer's trade deal (i.e., a temporary reduction in wholesale price). The result of this experiment suggests that retailer pass-through exceeds 100%. This finding is not necessarily inconsistent with profit maximization by the retailer. For example, monopoly pass-through is more than 100% if the demand function is log-convex (Amir, 2005).

There is an extensive marketing literature on estimation of consumer choice models using scanner data. Most of these studies find that brand switching plays a significantly more important role than in my study. Gupta's (1988) decomposition of the total price elasticity for regular ground coffee attributes 84% to brand switching. Chiang (1991) obtains similar results using a different choice model which treats the decisions of whether, what, and how much to buy simultaneously rather than independently. Both studies account for observed, but not unobserved, consumer characteristics. Chintagunta (1993) estimates a random-effects-discrete-choice model using panel data on yogurt purchases and obtains an elasticity decomposition in which 40% is due to brand switching, 15% is due to higher incidence, and 45% is due to switches to higher quantities. Bucklin, Gupta, and Siddarth (1998) also study the market for yogurt using a different approach to account for unobserved consumer heterogeneity and find that the percentages for the three effects are 58%, 20%, and 22% respectively. Bell, Chiang and Padmanabhan (1999) report price elasticity decompositions for 173 brands across 13 different product categories. They find that, on average, 75% of the elasticity is due to brand-switching, 11% to purchase incidence, and 14% to purchase quantity. However, Heerde, Gupta, and Wittink (2003) have recently pointed out that the way in which these studies measure the brand-switching effect is misleading and propose a complementary decomposition measure based on unit sales.[1] When they apply this measure to the same data used in Bell et al (1999), they find that brand switching accounts for only 33% of a sales bump. I adopt this unit sales

[1]In the elasticity-based decomposition, the brand-switching effect is measured as the change in brand choice probability conditional on purchase probability. In the unit-sales based decomposition, it is measured as the ratio of decreased sales of the non-promoted brands to increased sales of the promoted brand. The former measures the gross change in sales for non-promoted brands holding category volume constant whereas the latter measures the net change in sales of non-promoted brands, which takes into account the increase in category volume due to higher purchase probabilities.

approach for my decomposition exercise.

When I introduce wholesale prices from the Dominick's data into my analysis, one feature in particular stands out. Most price promotions are driven by the retailer, not the manufacturers. This is apparent from the considerable heterogeneity across stores in the frequency of price promotions. Some stores have frequent sales whereas others do not. The literature refers to the former as "high-low" stores and the latter as "everyday-low-price" stores. Using my demand estimates, I perform a counterfactual experiment in which I compare the observed price-cost margins with those that would be notionally optimal given the wholesale prices in the data. I find that regular brand prices at the "high-low" stores are typically at or below the optimal prices. By contrast, regular brand prices at the "everyday-low-price" stores are substantially below the optimal prices.

The immediate implication of these findings is that retailers seem to incur losses from their price promotions. Why then hold sales at all? The tendency of my model to over-predict retail markups and the variation in markups across stores suggest that retail price levels and price promotions reflect local competition among stores for customers who find shopping at different stores costly. Therefore, in a sense, Varian's (1980) theory of sales may have some relevance in explaining the existence of retailer price promotions: supermarkets use sales to compete for informed consumers (i.e., consumers who read supermarket flyers advertising products on sale). However, since stores sell many products and consumers usually choose stores based on the prices of multiple products in their shopping baskets, not just the price of one single item, incorporating retail competition into the current model is not straightforward and is not dealt with in this study.

There are several other theories of sales in the literature, but none appear to fit the present context well. Sobel (1984) develops an inter-temporal model of price discrimination in which a monopolist holds a sale from time to time in order to sell to low-valuation consumers who accumulate during periods of high prices. In his model, the product is a consumer durable and new consumers enter the market each period; neither feature characterizes the market for canned tuna. Aguirregabiria (1999) assumes that retailers stockpile products due to fixed costs of ordering and sales are needed to reduce heavy inventories. I lack data on inventories but the variation in sales across brands, stores, and time do not appear to be driven by inventory costs. Hendel and Nevo (2002, 2005) develop a demand model of consumer inventory holding and derive several implications for the observed data when stockpiling behavior is present.[2] Since canned tuna is a storable good, I initially considered their dynamic model of demand. Before estimating this model, however, I checked to see whether the data are consistent with the reduced-form predictions of Hendel and Nevo's model (2002). The results suggest that consumer stockpiling is not very important for this product category (see Chapter 3 for details). This is not too surprising, since most purchases of canned tuna consist of one or two cans.

I also consider other dynamic aspects of consumer behavior which might be thought to drive sellers' observed choices of sale depth and frequency. A

[2]Hong, McAfee, and Nayyar (2002) consider a model of price dispersion for a storable good, where price-sensitive consumers not only buy from the cheapest store, but can keep up to one unit in inventory, and sales are used to price discriminate between consumers with high and low search and storage costs. Anton and Das Varma (2005) consider a two-period Cournot model also for a storable good where they find that there exists an equilibrium such that prices are lower in the first period because firms have an incentive to steal rivals' demand from the future.

recent paper by Hartmann (2006) estimates a model of demand in which a store's customers have preferences for "spacing" in consumption. When I allow for such an effect in my demand estimates it turns out to be statistically significant, but very small in magnitude, too small to be a driving factor in sales patterns. Similarly, when I allow for a time-varying brand loyalty effect the estimated coefficient is small, negative, and insignificant. In summary, inter-temporal demand effects do not seem to explain the occurrence of sales in the data under examination here.

The rest of this chapter is organized as follows. Section 1.2 describes the data. The econometric model of demand is presented in Section 1.3 and the estimation procedure is described in Section 1.4. Section 1.5 reports the estimation results. Section 1.6 presents the estimated price elasticities and provides a unit-sales-based decomposition. Section 1.7 uses the demand estimates to compute retailer markups and pass-through rates. Section 1.8 investigates the profitability of retailer price promotions. Section 1.9 concludes.

1.2 Data

The consumer data I am using come from the ERIM data set provided by the James M. Kilts Center, GSB, University of Chicago. The ERIM data set was collected by the now-defunct ERIM division of A.C. Nielsen. It contains household-level purchase data for 9 packaged-good product categories in two mid-sized mid-western cities, Sioux Falls, SD and Springfield, MO, which are demographically typical of the U.S. population as a whole. Households who participated in the ERIM study were issued with magnetic ID cards. When they presented their ID cards at the checkout counters in the participating stores, their purchases of any UPC's in those 9 product categories were recorded. The total sales of the participating stores accounted for 80% of the market in terms of grocery and drug retail sales. Each record in the household-level data is a purchase occasion which records who made the purchase (household id), when (week) and where (store) the purchase occurred, the UPC and quantity being purchased, and the total payment made for the purchase. Moreover, when households joined the ERIM study, they were asked to fill out a questionnaire, which made available some basic demographics about the households. I use two demographic variables, household income and household size, to account for observed consumer heterogeneity in the model.

The demand model is estimated using purchase data for canned tuna in the city of Springfield, MO. The data consist of 58,920 purchase occasions of 5,255 households at 21 participating stores associated with 4 different supermarket chains. The sample period is from January 1985 to May 1987 (123 weeks). There are 5 package sizes in the canned-tuna product category: 3.25 oz, 6.5 oz, 9.25 oz, 12.5 oz, and 3-pack-3.25 oz. Table 1.1 presents the market share, purchase occasion share, and non-price promotion (i.e., feature advertisements or in-store displays) share of each package size. It shows that the 6.5 oz package size is the dominant one in terms of both store sales (89.6%) and household purchases (91.4%). Furthermore, Table 1.2 indicates that a handful of UPC's account for the majority of sales in this product category. They are 6.5-oz light tuna in water/oil of StarKist, Chicken of the Sea (CKN), Three Diamond, and private label (PL). The table shows the market, purchase-occasion, and non-price promotion shares, and average shelf prices of these 4 brands. They have a collective market share of more than 79% and account for at least 81% of the households' purchase occasions and 95% of the participating stores' non-price promotions. I only consider these top-selling UPC's in the demand estimation due to their dominance in this product category. Therefore, the number of consumers' brand choices in the demand model is limited to 4.

In the demand model, the alternatives in the choice set are the combinations of brands and units. To keep the model tractable, I limit the maximum quantity chosen to 4. It is not a very strong restriction since in 98.1% of the purchase occasions households purchased no more than 4 units of canned tuna. Moreover, when I define a sale as a price discount of at least 20% off the modal price, the percentage of sale purchase occasions such that the panel households purchased no more than 4 cans stays high (97.8%). Under these restrictions on brand choices and quantity choices, the households remaining in the sample are those who only purchased the stated brands and quantities, who had at least two purchase occasions, and whose demographic information is not missing. This selection leaves me 511 households and 5,143 purchase occasions.

In addition to the 16 "inside goods" (the combinations of 4 brand choices and 4 quantity choices), there is another alternative in the choice set, the "outside good", which represents the no-purchase occasion (of canned tuna) conditional on a store visit. Note that the no-purchase occasion is not directly observed in the data. To solve this problem, I construct a history of store visits for each household in the sample by collecting their purchase occasions in the 9 product categories in the ERIM data set. I then derive the no-purchase occasions for each household by comparing their store visit histories to their canned-tuna purchase histories. The drawback of this approach is that since the ERIM data set contains purchases in only 9 product categories, I cannot count those store visits where the household did not purchase any of the UPC's in those 9 product categories. An alternative way to derive the no-purchase occasions is to assume that a household visits a store every week, and for those weeks where store-visit data are missing, I can then assume that the household visited the store which appears most frequently in the "observed" store-visit history. The drawback of this approach is that even if a household does go grocery shopping weekly, for those inferred no-purchase occasions, the price and advertising information used to compute the choice probability in the likelihood function may not represent the true values of the data. I adopt the first approach and find 30,185 no-purchase occasions.

Retail prices of the 4 brands studied, particularly StarKist and CKN, exhibit a certain pattern over time: the prices remain constant at a "regular price" for a longer period of time and drop occasionally to a "sale price" for a shorter period. I find that prices are highly correlated across stores in the same chain. Therefore, I picked one store from each of the 4 supermarket chains in the data and graphed the weekly retail prices of StarKist and CKN in these 4 stores over the sample period, as shown in Figures 1.1 to 1.4. Even though we can easily observe the kind of price patterns just described, it is harder to pinpoint what the regular and sale prices are. I follow one approach in Hendel and Nevo (2002), by which regular price is defined as the modal price over the sample period. I then define a brand to be on sale if its shelf price is at least a certain percentage below its regular price. Table 1.3 presents the frequencies of sales of the 4 brands in each chain. It shows that frequencies of sales differ across brands within a chain. StarKist and CKN are promoted more often than the other two brands by the retailers. Moreover, frequencies of sales vary across chains within a brand. In addition, as illustrated in Figures 1.5 and 1.6, it seems that the timing and the magnitude of sales within a brand are not synchronized across chains. Another finding about these price promotions is that retailers seem to prefer to stagger sales of different brands across different weeks, rather than having more than one brand on sale simultaneously. Table 1.4 shows that in general, given a sale week, it is more likely that only one brand is on sale.

1.3 Demand Model

Consumers' purchase decisions on whether to buy (purchase incidence), what to buy (brand choice), and how much to buy (quantity choice) are modeled jointly using a nested logit model. Let i denote consumer ($i = 1, \ldots, N$), j denote brand choice ($j = 1, \ldots, J$), x denote quantity choice ($x = 1, \ldots, X$), and t denote time of a store visit (T_i is the total number of store visits by consumer i). Given that consumer i visits a store at week t, the consumer problem is to choose one of the alternatives in the choice set $\{0, jx; \forall j \forall x\}$ (where 0 denotes the choice of outside good and jx denotes the choice of x units of brand j) to maximize utility. The utility derived from the outside good is

$$u_{i0t} = \gamma_{i0} + \beta_{i1} H_{it} + \beta_{i2} H_{it}^2 + \zeta_{i0t} + \rho_i \varepsilon_{i0t} \tag{1.1}$$

The utility derived from the inside good jx is

$$u_{ijxt} = \gamma_{ij} x^{\alpha_i} + \gamma_{iP} P_{jt} x + \gamma_{iA} A_{jt} + \zeta_{ixt} + \rho_i \varepsilon_{ijxt}, \tag{1.2}$$

$$\text{for } j = 1, \ldots, \text{J}, \; x = 1, \ldots, X$$

Here γ_{i0} and γ_{ij}'s are consumer i's taste preferences on the outside good and on the brands of inside goods. α_i is a measure of how fast marginal utility decreases with respect to quantity consumed and is constrained to $0 < \alpha_i \leqslant 1$. H_{it} is the elapsed time since last consumption (i.e., the interval between consumer i's last purchase occasion and the current shopping occasion, measured in weeks), and β_{i1} and β_{i2} are the parameters that determine how consumer i's utility is affected by H_{it}. P_{jt} is the price of brand j at week t, and γ_{iP} is consumer i's marginal utility of income. A_{jt} is a dummy variable which is 1 if brand j is on feature/display advertising at week t and 0 otherwise, and γ_{iA} measures consumer i's response to advertising. $\zeta_{igt}(g = 0, 1, \ldots, X$, X=4) is an unobserved shock common to every alternative in the group g. $\varepsilon_{ikt}, k \in \{0, jx, \forall j \forall x\}$ is another unobserved shock assumed identically and independently distributed extreme value. The distribution of ζ_{igt} is the unique distribution with the property such that if ε_{ikt} is an extreme value random variable, the total unobserved utility, $\zeta_{igt} + \rho_i \varepsilon_{ikt}$, is also an extreme value random variable. Finally, ρ_i measures the degree of independence in unobserved utility among the alternatives within a group. It is assumed to be individual-specific but the same across groups and its value is constrained to $0 < \rho_i \leqslant 1$.

In the standard discrete-choice logit model, the error term is used to account for the unobserved components of utility and is assumed to be i.i.d. across alternatives and time. However, if the observed product characteristics do not account for all the important aspects of the choice set, the unobserved utilities of some alternatives may be correlated with each other due to some unaccounted dimensions in the product characteristics space. I use brand dummies as the observed product characteristics - these dummies are then subsumed into the γ_{ij}'s. To account for the possibility of unobserved shocks to consumption quantities, I nest alternatives with the same quantity of different brands into the same group. This specification captures the possibility that when a consumer has a high unobserved demand shock for the alternative, say, x units of brand j, his/her unobserved demand shocks for x units of other brands may be high as well. A higher value of ρ_i means greater independence and less correlation in the unobserved components of utility among the alternatives within a nest. In the polar case such that ρ_i equals 1, the nested logit model reduces to the standard logit model (Train, 2003, Chapter 4).

I assume that households do not stockpile so that a purchase occasion coincides with a consumption occasion. At first glance, this assumption may seem too strong to be applied to the canned-tuna product category. Since canned food is storable, consumers can stock up when it is on sale. However, storability is only a necessary condition for stockpiling. When a good is not a necessity and has a low consumption frequency, price promotions of the good may only prompt consumers to purchase for current consumption, rather than to stockpile for future consumption since they may incur an inventory cost. Chapter 3 examines whether consumer stockpile canned tuna during sales. Using the same data set, I check whether the data are consistent with the reduced-form predictions of Hendel and Nevo's model of consumer inventory holding (2002). The results suggest that only a limited amount of stockpiling is present. Therefore, we need not to be too concerned about potential bias in the price elasticities from the current static model with no consumer stockpiling.

When the product category under consideration is food and a non-necessity, consumers' purchase behavior may be influenced by the passage of time. To account for this possibility, I allow a consumer's utility from the outside good to vary over time depending on the elapsed time since last consumption (purchase). Therefore, the elapsed time since last consumption only affects the probability of purchase, not the probability of brand-quantity choice conditional on a purchase. If this variable has a positive impact on the purchase probability for the outside good, then it suggests a model of preferences in which desire for tuna declines over time, i.e. the longer the time period without consuming canned tuna, the less likely a purchase will be made. On the other hand, if the impact is negative, it suggests a model of preferences in which desire for canned-tuna increases over time, i.e., utility from consumption is increasing with respect to the time since last consumption (Hartmann, 2006).

The problem of consumer utility maximization in the nested logit model has a closed-form solution. First of all, the probability that consumer i chooses the alternative of x units of brand j at week t is

$$\begin{aligned} Pr_{it}(jx) &= Pr_{it}(j|x) \times Pr_{it}(x) \\ &= \frac{\exp(\delta_{ijxt}/\rho_i)}{\sum_{k=1}^{J} \exp(\delta_{ikxt}/\rho_i)} \times \frac{(\omega_{ixt})^{\rho_i}}{1 + \sum_{y=1}^{X} (\omega_{iyt})^{\rho_i}} \end{aligned} \tag{1.3}$$

where

$$\delta_{ijxt} = \gamma_{ij} x^{\alpha_i} + \gamma_{iP} P_{jt} x + \gamma_{iA} A_{jt} - (\gamma_{i0} + \beta_{i1} H_{it} + \beta_{i2} H_{it}^2),$$

$$\text{for } j = 1, \ldots, J, \; x = 1, \ldots, X$$

$$\omega_{ixt} = \sum_{j=1}^{J} \exp(\delta_{ijxt}/\rho_i), \text{ for } x = 1, \ldots, \text{X}$$

The probability that consumer i chooses the outside good, i.e., no purchase of canned tuna, at week t is

$$Pr_{it}(0) = \frac{1}{1 + \sum_{y=1}^{X} (\omega_{iyt})^{\rho_i}} \tag{1.4}$$

Then, the probability of purchase incidence can be computed by

$$1 - Pr_{it}(0) = \frac{\sum_{x=1}^{X} (\omega_{ixt})^{\rho_i}}{1 + \sum_{x=1}^{X} (\omega_{ixt})^{\rho_i}}$$

Finally, market demand for brand j can be computed by aggregating demand across quantities and across consumers, i.e.,

$$D_j(p) = \sum_{i=1}^{N} \sum_{x=1}^{X} x Pr_{it}(jx), \qquad \text{for } j = 1, \ldots, J \tag{1.5}$$

To account for consumer heterogeneity, all the parameters in the model are individual-specific. However, since the length of the panel data is not long, I assume that these individual-level parameters do not vary over time. Consumer heterogeneity is modeled by a random-coefficients approach where heterogeneity is driven by observed consumer demographics and an unobserved component, interpreted as unobserved consumer characteristics. I assume that the distribution of individual-level parameters follows a conditional multivariate normal with mean as a function of consumer demographics, i.e.,

$$\theta_i = \Pi D_i + \nu_i, \tag{1.6}$$

$$\text{where } \nu_i \sim iidN(0, diag(\Sigma_\theta)), \text{ for i} = 1, \ldots, N$$

Here θ_i is a $k \times 1$ vector consisting of k individual-level parameters, i.e.,

$$\theta_i = (\{\gamma_{ij}\}_{j=1}^{J}, \gamma_{i0}, \gamma_{iP}, \gamma_{iA,}\beta_{i1}, \beta_{i2}, \alpha_i^*, \rho_i^*)'$$

D_i is a $d \times 1$ vector consisting of a constant and $d-1$ demographic variables; Π is a $k \times d$ coefficient matrix which measures how individual-level parameters vary with demographics; Σ_θ is a $k \times 1$ coefficient vector which determines the dispersion of unobserved consumer characteristics. Rossi, McCulloch, and Allenby (1996) suggest that the variability of individual-level parameters due to unobserved rather than observable consumer heterogeneity can be measured by comparing Σ_θ with the marginal variance of θ_i. Finally, because of the constraints on the values of α_i and ρ_i, I define

$$\alpha_i = \frac{\exp(\alpha_i^*)}{1 + \exp(\alpha_i^*)} \text{ and } \rho_i = \frac{\exp(\rho_i^*)}{1 + \exp(\rho_i^*)}.$$

That is, α_i and ρ_i are transformations of normals and the constraints are always satisfied.

Consumer demand is modeled by a discrete-choice-logit model where the choice set consists of all combinations between the brand and quantity choices. The reason for this approach is that in the data households usually purchase one or more units of only one brand on a given shopping trip. However, if it is the case that consumers purchase multi-units as well as multi-brands at the same time, other approaches will have to be employed to handle such multiple discreteness. One example is the approach proposed by Hendel (1999) and Dube (2005). Hendel studies the personal computer (PC) market, and in his data, firms purchase multiple units of different brands of PCs. Dube studies the carbonated soft drink (CSD) industry, and in his CSD purchase data, multiple-item purchases occur frequently. In Hendel's model, demand for assortments results from the number of potential tasks each firm has that can be performed by PC's. In Dube's model, the demand for assortments reflects the number of potential consumption occasions a household expects before the next shopping trip. The Hendel/Dube approach is to derive a purchase demand system by aggregating over unobserved consumption occasions. It is more powerful in the sense that it is able to deal with multiple discreteness. However, since households typically purchase only one brand at each shopping trip in my data, their model does not fit the data here better than the logit model. That is, the main difference between my data and standard single-unit purchase data is in multiple-unit purchases of a given brand, which I allow for by treating quantity choices as the nests in my logit.

1.4 Estimation Procedure

I estimate the model using the Markov Chain Monte Carlo (MCMC) method, which is a simulated Bayesian method and is well-suited for models built from a sequence of conditional distributions (Rossi and Allenby, 2003). Compared to classical procedures such as maximum likelihood methods, Bayesian procedures do not require the maximization of any function. Moreover, desirable estimation properties such as consistency and efficiency can be attained under more relaxed conditions. However, since parameters in the Bayesian procedures are represented by a posterior distribution which reflects both prior beliefs and sample information, to derive relevant Bayesian statistics requires an iterative process which takes draws from the posterior and converges after a sufficient number of iterations. The problem is that in practice it is not easy for researchers to determine when convergence has been achieved. Therefore, even though Bayesian procedures circumvent the difficulties of convergence to a maximum in the classical approaches, they have their own convergence problems. However, the way to examine and interpret an estimator derived by Bayesian procedures is no different from that in classic procedures. In fact, under certain conditions, the Bayesian estimator is asymptotically equivalent to the maximum likelihood estimator. Therefore, Train suggests that a researcher's statistical perspective should not dictate which estimation procedure to adopt. Instead, what the researcher should consider is which type of convergence will be less burdensome regarding the particular setting of her/his model (Train, 2003, Chapter 12).

Consumer heterogeneity is modeled by a random-coefficients approach where heterogeneity is driven by observed household demographics and unobserved components with a known distribution. Therefore, the demand model can be specified through a series of conditional distributions, also known as a hierarchical model, which the MCMC method is especially suited to estimate (Allenby and Rossi, 1999). Moreover, the MCMC method enables inference not only about the common random-coefficients parameters, but also about the individual-specific parameters. Note that any successful marketing analysis requires a detailed understanding of the distribution of consumer heterogeneity. However, when the preferences of the consumers targeted by marketers are not well represented by the summary statistics describing the distribution of heterogeneity for the entire market, the estimated responses obtained using only the aggregate information can be biased. Under these circumstances, inferences about the unit-level parameters enabled by the the MCMC method become especially useful since the market-wide response can be derived by aggregating over the individual household's response to make the adjustments for the distribution of heterogeneity (Rossi and Allenby, 1993).

Roughly speaking, the MCMC estimation method proceeds by specifying a prior distribution for the parameters in the model and then repeatedly taking draws from the posterior distribution conditional on the observed data, which can be facilitated by Gibbs sampling and/or the Metropolis-Hasting (MH) algorithm. Bayes theorem gives the following relation.

$$P(\Omega|Y) \propto L(Y|\Omega)P(\Omega) \tag{1.7}$$

Here Ω is the vector of model parameters, including the population-level parameters (Π and Σ_θ) and the individual-level parameters (θ_i's). $P(\Omega)$ is the prior distribution of Ω. Y is the set of observed choices made by independent decision makers in the sample. $L(Y|\Omega)$ is the likelihood function of the observed choices which depends on Ω. And $P(\Omega|Y)$ is the posterior distribution of Ω, which is proportional to the prior distribution multiplied by the likelihood function, reflecting both the prior beliefs and the sample information.

The hierarchical model in this application consists of the consumer-level likelihood, $L(y_i|\theta_i), \forall i$, the first-stage prior, $\phi(\theta_i|\Pi, \Sigma_\theta), \forall i$, and the second-stage prior, $K(\Pi, \Sigma_\theta|\Lambda)$. Since households' choices are modeled by nested logit, the unit-level likelihood function is a product of multinomial logits, i.e.,

$$L(y_i|\theta_i) = \prod_{t=1}^{T_i} Pr(y_{it})$$

where

$$Pr(y_{it} = jx) = \frac{\exp(\delta_{ijxt}/\rho_i)}{\sum_{k=1}^{J} \exp(\delta_{ikxt}/\rho_i)} \times \frac{(\omega_{ixt})^{\rho_i}}{1 + \sum_{y=1}^{X} (\omega_{iyt})^{\rho_i}},$$
$$\text{for } j = 1, \ldots, 4,\ x = 1, \ldots, 4$$
$$Pr(y_{it} = 0) = \frac{1}{1 + \sum_{y=1}^{X} (\omega_{iyt})^{\rho_i}}$$

Here $\phi(\theta_i|\Pi, \Sigma_\theta)$ denotes the normal density since the distribution on individual-level parameters is specified as a conditional multivariate normal, i.e., $\theta_i \sim N(\Pi D_i, \Sigma_\theta)$, for $i = 1, \ldots, 511$. $K(\Pi, \Sigma_\theta|\Lambda)$ is the prior on the population-level parameters, Π and Σ_θ, and Λ is the parameters of the prior distribution. I assume that the prior on Π (denoted by $k(\Pi)$) is normal with mean μ_0 and variance matrix σ_0^2, i.e., $\Pi \sim N(\mu_0, \sigma_0^2)$; the prior on Σ_θ (denoted by $k(\Sigma_\theta)$) is inverted Wishart with ν_0 degrees of freedom and a scale matrix s_0, i.e., $\Sigma_\theta \sim IW(\nu_0, s_0)$; and the two priors are independent. Moreover, I use very diffuse priors on Π and Σ_θ by setting

$$\mu_0 = 0,\ \sigma_0^2 = 100I,\ \nu_0 = \dim(\Sigma_\theta) + 1, \text{ and } s_0 = I.$$

The values of the prior parameters are chosen to make the two priors nearly flat, i.e., all possible values of the parameters are considered equally likely.

Therefore, the prior distributions represent very little knowledge about the parameters before taking the sample and the sample information about each household greatly influences the posterior.

The joint posterior distribution of the model parameter is described by Equation (1.8).

$$P(\Pi, \Sigma_\theta, \theta_i \forall i | Y) \propto \prod_{i=1}^{N} \prod_{t=1}^{T_i} \Pr(y_{it}|\theta_i) \phi(\theta_i|\Pi, \Sigma_\theta) \times k(\Pi) k(\Sigma_\theta) \tag{1.8}$$

Taking draws from this posterior can be done by Gibbs sampling, which involves breaking the model parameters into different blocks and successively taking one draw in each block in turn, conditional on the values of the parameters in the other blocks. I break the model parameters into 3 blocks: the first block contains the individual-level parameters, θ_i's, the second and third blocks contain the population-level parameters, Π and Σ_θ. I take draws of the model parameters from the posterior through a three-step procedure described as follows.

Step 1: Take a draw of θ_i conditional on the values of Π and Σ_θ, for each household in the sample. The conditional posterior distribution of θ_i is

$$P(\theta_i|\Pi, \Sigma_\theta, y_i) \propto \prod_{t=1}^{T_i} \Pr(y_{it}|\theta_i) \times \phi(\theta_i|\Pi, \Sigma_\theta), \text{ for } i = 1, ..., 511$$

There is no simple way to draw from this conditional posterior because it involves multinomial logits. Thus the Metropolis-Hastings algorithm is used to take draws in this step.

Step 2: Take a draw of Π conditional on the values of Σ_θ and θ_i's. The conditional posterior distribution of Π is

$$P(\Pi|\Sigma_\theta, \theta_i \forall i) \propto \prod_{i=1}^{N} \phi(\theta_i|\Pi, \Sigma_\theta) \times k(\Pi)$$

This conditional posterior follows a normal distribution. Thus taking draws in this step is straightforward.

Step 3: Take a draw of Σ_θ conditional on the values of Π and θ_i's. The conditional posterior distribution of Σ_θ is

$$P(\Sigma_\theta|\Pi, \theta_i \forall i) \propto \prod_{i=1}^{N} \phi(\theta_i|\Pi, \Sigma_\theta) \times k(\Sigma_\theta)$$

This conditional posterior follows an inverted Wishart distribution. Thus taking draws in this step is also straightforward. Details on how draws are taken from these conditional posteriors are provided in the appendix.

This three-step procedure is repeated to obtain a sequence of draws, which after enough iterations converges to draws from the joint posterior of the model parameters. The initial draws in the sequence prior to convergence, often called burn-in, should be discarded. Moreover, since each iteration obtained through Gibbs sampling builds on the previous one, these draws will be correlated over iterations even after convergence has been achieved. To reduce the amount of correlation among the draws, Train (2003) suggests using only a portion of the draws that are obtained after convergence. 70,000 iterations of the Gibbs sampling are performed in my estimation. I discard the initial 50,000 iterations and retain every tenth draw in the subsequent 20,000 iterations. I then use the retained 2,000 draws to compute the estimates and the standard errors of the model parameters.

1.5 Estimation Results

Table 1.5 presents the MCMC estimation results for the population-level parameters. The first 3 columns are the estimates and standard errors of Π and the fourth column Σ_θ. Based on the estimates of households' taste parameters for brands of the inside goods and the outside good, in general, consumers prefer the outside good the most and then StarKist. This is reflected in the high frequency of no purchase and the high market share of StarKist in the data. The estimates of the interaction terms between brands and demographics indicate that household income has the highest impact, and household size has the lowest impact on the outside good, but they are not significant. The estimates of households' responses to price and to feature/display advertising are as expected, i.e., negative and positive, respectively. The estimates of the interaction terms between price/advertising and demographics indicate that households with higher income are less sensitive to both price and feature/display and larger households are more price sensitive but less advertising sensitive. However, the estimates of the interaction terms involving household size are not significant.

In general, the utility derived from the outside good is concave with respect to the elapsed time since last consumption, i.e., the effect of the lapse of consumption on the no purchase utility is positive first and then turns negative. For example, for households with median income ($27,500) and median size (3 members), Figure 1.7 illustrates how their utility from the outside good changes w.r.t. the elapsed time since last consumption. It seems that households' decisions of whether or not to consume canned tuna initially exhibit a degree of habit persistence, i.e., over some interval a category purchase becomes less likely with the passage of time. However, after a sufficiently long time without consuming canned tuna, households eventually become more likely to make a purchase as time passes. Chiang (1991) includes the length of the interval since last purchase of any brand in his model as well. His estimates suggest a negative impact of time on the no purchase probability, which is as expected since he uses this variable as an inventory proxy. That is, a longer time since last purchase is interpreted to mean lower inventories, and thus a higher likelihood of making a purchase during the current shopping trip. In my model, I assume that households do not stockpile, so a negative relationship between time-since-last-purchase and the no purchase probability is interpreted as a changing preference for consuming canned tuna. Note that if the no-stockpiling assumption did not hold, it would be difficult to separate stockpiling behavior from preferences for spreading. Based on the estimation results, in general, the impact of the elapsed time since last consumption on the no purchase probability is positive when the lapse is no longer than 26 weeks, which, in a way may be used to substantiate the earlier finding that only a limited amount of stockpiling is present in this product category. Moreover, even though the estimated coefficient for this variable is statistically significant, the magnitude is very small, which makes little differences in terms of aggregate demand over time. An exercise is performed to investigate its effect on category sales over time following a price promotion. I find that if there is a sale (of any brand and a price discount of 10% to 50%) in a week, on average, category sales in the following non-sale week increase only 0.37% and the magnitude of this change becomes smaller and smaller over time. For example, after 5 weeks, it is less than 0.1%. Average sales of canned tuna in a busy store during a non-sale week are about 400 cans, which means that the increased sales are less than 2 cans the week after a price promotion. Since the inter-temporal demand effect caused by the elapsed time since last consumption is so small, retailers or manufacturers probably do not consider this variable when making their pricing decisions. I thus ignore it when simulating consumer demand in the later exercises.

Recall that the coefficients, α_i and ρ_i, are transformations of normals. As an example, for a household with median income and size, α_i is about 0.110 and ρ_i is about 0.496.[3] The results imply that marginal utility decreases at a pretty fast rate with respect to quantity consumed and the unobserved components of utility for the alternatives in the same nest (i.e., alternatives of same quantity across brands) are indeed correlated.

Finally, the nested-logit-random-coefficients model of consumer demand can be estimated by either the simulated maximum likelihood method or the MCMC methods. However, the latter enables the recovery of individual-level parameters. The value of this additional information depends on how much consumer heterogeneity is accounted for by observed consumer characteristics. Table 1.6 provides such an evaluation. Note that Σ_θ measures the conditional variance of individual-level parameters, θ_i's (i.e., the dispersion of θ_i across households conditional on demographics). Therefore, by combining the magnitude of Σ_θ with the unconditional variability of the elements in θ_i across households, we can estimate the proportion of heterogeneity which is accounted for by the observed consumer characteristics (Rossi, McCulloch, and Allenby, 1996). I define a statistic, λ^2, as $\lambda^2 = 1 - \frac{var(\nu)}{var(\theta)}$, where $var(\nu)$ is the variance of the unobservable component and $var(\theta)$ is the total variation. λ^2 measures the proportion of the variation in the values of individual-level parameters which can be explained by household demographics. According to Table 1.6, the demographic variables are best at explaining the variation in households' parameters on CKN (70%), StarKist(67%), and the outside good

[3]These individual-level parameters are calculated based on the estimates reported in Table 1.5. $\alpha_i = \frac{\exp[-1.934+(-0.127)\times\ln(2.75)+(-0.011)\times 3]}{1+\exp[-1.934+(-0.127)\times\ln(2.75)+(-0.011)\times 3]}$ and $\rho_i = \frac{\exp[-0.011+(-0.123)\times\ln(2.75)+0.071\times 3]}{1+\exp[-0.011+(-0.123)\times\ln(2.75)+-0.071\times 3]}$.

(57%), are modest at explaining the variation in households' parameters on the private label (37%), Three Diamond (34%), and the elapsed time since last consumption (36% and 30%), are not so good at explaining the variation in price sensitivity (17%), and are worst at explaining the variation in the values of the rest (less than 7%). I think the inference about individual-level parameters enabled by the MCMC method is useful here since demographic variables have limited value in predicting how households respond to prices, while price sensitivity is indeed a key determinant for price elasticities and thus has important implications for potential profitability of price promotions. As a result, market-wide responses to price promotions and own- and cross-price elasticities are estimated by aggregating across the responses of the sample households, utilizing the recovered individual-level parameters to account for parameter uncertainty.

1.6 Price Elasticities and Sales Bump Decomposition

Table 1.7 reports the own- and cross-price elasticities of the 4 brands and the elasticities of the no purchase probability. Price elasticities are computed by aggregating over the simulated responses of the sample households using the estimated individual-level parameters. I compute the elasticity matrices for each store-week when the prices are distributed as in the data and take the average.

The orders of the own-price and no purchase elasticities are close to the order of the market shares of the 4 brands, e.g., StarKist which has the highest market share is least price elastic and is most likely to cause category expansion when its price is dropped. Note that the estimated brand-preference coefficient for StarKist is the highest compared to those for the other 3 brands.

Therefore, assuming that the magnitude of brand-preference coefficients is an indicator of brand quality, it seems that a price promotion on the brand with higher quality has a greater category expansion effect. The low variation in the cross-price elasticities in each column may suggest the independence-of-irrelevant-alternative property of a logit model. However, it is not easy to eliminate this problem since the products included in the demand system are pretty homogenous and I could not think of other product characteristics, except for "brand", to differentiate these products.

As often observed in the data, a temporary price reduction can boost the quantity demanded for a brand and thus generate a sales bump. Increased demand for a brand that is on sale may stem from three different responses to price promotions: first, some consumers who are buying other brands when brand X is not on sale may switch to brand X when brand X is on sale; second, some consumers who are buying brand X at its regular (non-sale) price may purchase a greater quantity of brand X when it is on sale; and third, some consumers who buy nothing in this product category without any price promotion may decide to purchase brand X when it is on sale. Therefore, a promotional sales bump can be decomposed into three components: brand-switching, purchase quantity, and purchase incidence.

Marketing researchers are interested in the relative sizes of these three components because of their implications for the relative profitability of price promotions as perceived by manufacturers and retailers. If a sales bump comes mainly from consumer brand switching, then it profits one manufacturer (of brand X) at the expense of another, and may be of relatively little benefit to the retailer (if, e.g., the margin on brand X is lower when it is on sale). Retailers may profit more if price promotions have a large impact on purchase incidence and purchase quantity, since overall sales of the whole category then increase. To study the retailer-manufacturer split of the benefits from price promotions, I simulate the sample households' reactions to price promotions using the demand-side estimates and decompose a sales bump into the brand-switching, purchase-incidence, and purchase-quantity effects.

Table 1.8 presents the results of sales bump decomposition with different magnitudes of retail price discount and the percentage increases in sales of the promoted brand and the product category as a whole. It is computed as follows. First, I assume the "regular" prices of the 4 brands to be the average shelf prices over the sample period (shown in the last column of Table 1.2), and I compute the "baseline market demand" for each brand at their regular prices by aggregating across the demand of the sample households. Second, for each brand in turn, I drop its regular price by a certain percentage (10%-50%), and I decompose the sample households' responses to price promotions into 3 components: brand switching, purchase incidence, and purchase quantity by comparing the changes in their choice probabilities due to these price reductions. Finally, the market-wide decomposition is derived by aggregating across the sample households' responses.

As the magnitude of discount changes, the proportions of the increased sales of the promoted brand which can be attributed to brand switching, purchase incidence, and purchase quantity change too. In general, the greater is a price discount, the lower is the proportion attributable to brand switching and the higher are the proportions attributable to purchase incidence and purchase quantity (except for StarKist). Comparing the results among the 4 brands, it seems that the higher the quality of the promoted brand as measured by the estimated brand-preference coefficients, the higher is the proportion of increased sales attributable to purchase incidence, and the lower the quality of the promoted brand, the higher is the proportion of the increased sales attributable to brand switching. Moreover, the percentage increase in sales of the promoted brand is higher for lower-quality brands and the percentage increase in sales of the category as a whole is greater for higher-quality brands. Therefore, price promotions on the brand with the highest quality have the greatest category expansion effect, which is also implied by the elasticities of no purchase probability reported in Table 1.7. The generally large proportion of increased sales attributable to purchase incidence implies that in the canned tuna product category price promotions are relatively more effective in moving consumers from no purchase to purchase.

Most papers in the marketing leterature which study consumers' purchase behavior and decompose consumers' responses to marketing variables use are elasticity-based decompositions. For example, Gupta (1988), Chiang (1991), Chintagunta (1993), Bucklin, Gupta, and Siddarth (1998), and Bell, Chiang and Padmanabhan (1999) model all three purchase decisions simultaneously (i.e., when, what and how much to buy), so the total price elasticity can be obtained as a sum of brand-choice elasticity, purchase-incidence elasticity, and purchase-quantity elasticity. Across these studies, on average 74% of the total price elasticity is attributable to brand switching. Intuitively, one may tend to interpret this result as indicating that if the increased sales of the promoted brand are 100 units, the sales of non-promoted brands in the same product category decrease by 74 units. Then price promotions look to be more effective in drawing consumers from competitive brands than in expanding category sales. However, Heerde, Gupta, and Wittink (2003) point out that such an interpretation, i.e., interpreting the percentage of elasticity due to brand switching as the ratio of the lost sales of the non-promoted brands to the increased sales of the promoted brand, is incorrect. This is because the brand-switching effect in the elasticity-based decomposition measures the changes in brand choice probability conditional on purchase occurrence, i.e., it measures the gross change in sales for the non-promoted brands when category volume is held constant. However, since price promotions usually have a positive effect on purchase-incidence probability, the part of the increased sales of the promoted brand that is attributable to the decreased sales of the non-promoted brands (i.e., the net change) will be smaller when the increase in category volume is accounted for.

Heerde, Gupta, and Wittink (2003) thus propose a complementary decomposition measure which is based on unit sales and apply this measure to the same data used in Bell, Chiang, and Padmanabhan (1999). They find that the 75% brand-switching effect reported by Bell et al. translates to only 33% using the unit sales decomposition. The decomposition results reported in Table 1.8 are unit sales based and they show that in the canned tuna product category, the sales bump attributable to brand switching is about 36% on average across brands and magnitudes of discount. Therefore, the unit-sales-based decomposition results suggest that the brand-switching effect of price promotions may be much smaller than has been previously assumed, and may be more consistent with the frequent sales observed in retail stores.

1.7 Retailer Markups

In this section, the demand estimates are applied to the supply side to investigate the pricing behavior of retailers. Retailer r's operating profit at week t is given by

$$\Pi_{rt}(p_{rt}) = \sum_{j=1}^{J}(p_{rjt} - wp_{rjt})q_{rjt}(p_{rt}) \tag{1.9}$$

Here p_{rjt} and wp_{rjt} are retailer r's price and marginal cost (i.e., wholesale price) of product j at week t respectively, and q_{rjt} is the demand for product j retailer r faces at week t. If we assume that retailers choose retail prices to maximize per period category-level profits, then the profit-maximizing retail prices, p_{rt}, should satisfy a system of first order conditions described by Equation (1.10).

$$q_{rt}(p_{rt}) + \Omega_{rt}(p_{rt} - wp_{rt}) = 0 \tag{1.10}$$

where

$$q_{rt} = \begin{pmatrix} q_{r1t} \\ \vdots \\ q_{rJt} \end{pmatrix}, \ \Omega_{rt} = \begin{pmatrix} \dfrac{\partial q_{r1t}}{\partial p_{r1t}} & \cdots & \dfrac{\partial q_{rJt}}{\partial p_{r1t}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial q_{r1t}}{\partial p_{rJt}} & \cdots & \dfrac{\partial q_{rJt}}{\partial p_{rJt}} \end{pmatrix}, \ p_{rt} - wp_{rt} = \begin{pmatrix} p_{r1t} - wp_{r1t} \\ \vdots \\ p_{rJt} - wp_{rJt} \end{pmatrix}$$

Therefore, retailer r's marginal costs (and markups) can be backed out using Equation (1.11).

$$wp_{rt} = p_{rt} + \Omega_{rt}^{-1} q_{rt}(p_{rt}) \tag{1.11}$$

Table 1.9 reports the average marginal costs (across time and stores within the same chain) for each brand-chain. It shows that not only do marginal costs vary across brands carried by the same retailer, but the marginal costs of the same brand vary across retailers as well. A possible explanation for the former is that manufacturers may have different production costs or markups. As regards the latter, retailers have different degrees of bargaining power against manufacturers in negotiating wholesale prices, or they may incur some sort of retailer-specific costs.

The estimates reported in Table 1.9 implicitly assume that price promotions are induced by manufacturers' trade deals, i.e., manufacturers sell their products at a temporary price reduction to retailers. If that is the case, we can estimate retailer pass-through, defined as the ratio of retail price reduction to wholesale price reduction. I conduct a counterfactual experiment to measure the possible range of retailer pass-through in the canned-tuna product category. For simplicity, I focus on only two brands, StarKist and CKN, since the manufacturers of these two brands are well-established big companies and are more likely to be able to afford and offer trade deals to retailers. I fix the non-sale retail prices of the two brands at \$0.89 (which is the modal price observed in the data) and vary the sale prices to be \$0.69, \$0.59, and \$0.49. 4 scenarios are considered: no sale for both StarKist and CKN, a sale for both, a sale only for StarKist, and a sale only for CKN. I then use Equation (1.11) to back out the associated wholesale prices, assuming that the retailer's marginal costs consist of only wholesale prices.

Table 1.10 presents the results of this counterfactual experiment. It shows that given a trade deal provided by the manufacturer, the retailer passes on to consumers the full amount of the wholesale price reduction, and then some, i.e., retailer pass-through exceeds 100%. Moreover, retailer pass-through is higher when sales of the two brands do not occur at the same time, and it decreases with respect to depths of the deals. The empirical findings from prior research on manufacturers' trade deals show that retailer pass-through can range from 0% to more than 100% (Tyagi, 1999). A pass-through rate of more than 100% may seem counter-intuitive at first glance, but it is consistent with a log-convex demand function (Amir, 2005). Since I do not have a closed-form demand function and retailers may be more concerned about category sales, I simulate category demand by changing StarKist's (CKN's) retail price while fixing CKN's (StarKist's) price at the non-sale level ($0.89). Figure 1.8 displays the logarithm of these two simulated category demand functions, and they are convex.

The analysis of markups and pass-through assumes that demand is static, which means the retailer does not face any inter-temporal trade-offs in pricing its products. As mentioned earlier, a reduced-form examination suggests that consumer stockpiling is not an important factor in this product category (see Chapter 3). Although I have taken into account the possibility that preferences may change over time by allowing utility to depend on the elapsed time since last consumption, the magnitude of this effect on quantity demanded is very small and can be ignored (see Section 1.5). I also considered some form of time-varying brand loyalty by allowing utility to depend on the brand consumers purchased last time. The estimated coefficient for this variable turns out to be small, negative, and insignificant. Therefore, it seems that retailers have no reason to engage in inter-temporal price discrimination nor do they face any significant dynamics in consumer demand that may arise from non-additive preferences.

The analysis also ignores any dynamics that may arise from retailer inventory decisions. This assumption is more problematic. Retailers may take advantage of manufacturers' trade deals and forward buy or they may incur fixed ordering costs, both of which would make retailers keep inventory and thus influence their pricing behavior. See Dreze and Bell (2001) for more details on the former issue and Aguirregabiria (1999) on the latter.

1.8 Retailer Price Promotions

A more fundamental objection to the analysis presented in the preceding section is that retailers may not be choosing prices to maximize category-level profits. Since supermarkets sell many products, the price level of a specific product could affect purchases not only of that product category, but also of other product categories. Therefore, retailers may be coordinating their pricing decisions across different product categories to build up store traffic and maximize store-level profits. Indeed, some of the price reductions observed in the data are likely to represent retailer sales, not manufacturer sales. The simple model of pricing given above cannot explain price promotions by retailers.

To investigate this issue, I resort to another data set from Dominick's Finer Foods (DFF) where a proxy for wholesale prices is available which helps to distinguish price promotions by retailers from price promotions by manufacturers. The DFF database is provided by the James M. Kilts Center, GSB, University of Chicago. Dominick's is the second biggest supermarket chain in the metropolitan area of Chicago, Illinois. It has about 100 stores and accounts for approximately 25% of the market in this area (Chevalier, Kashyap and Rossi, 2000). The DFF database gives weekly retail prices, unit sales (quantities), and retail profits for each UPC in 29 different product categories (including canned tuna) in each store from September 1989 to May 1997. Defining retail profits as percentage gross margins, gross costs can then be recovered by multiplying retail prices with retail profits. However, these costs represent average acquisition costs (AAC's) of the items in inventory and they do not exactly correspond to wholesale prices. For example, a temporary wholesale price cut this period may take time to work itself into AAC's as the old higher-priced inventory carried by the retailer is being sold off. In this case, wholesale prices will be smaller than AAC's. Alternatively, if the retailer engages in forward buying during a trade deal, then the AAC's will remain low for some time after the wholesale prices have gone back up (Chevalier et al, 2000). In this case, wholesale price will be greater than AAC's. Despite these measurement errors, manufacturer-driven price promotions are still observable since off-invoice discounts (which are the most important component of manufacturers' trade promotion expenditures) and promotional moneys for feature advertisements and special displays are accounted for and thus reflected in AAC's (Chintagunta, 2002). There may exist other promotional agreements between the manufacturers and Dominick's which do not show up in AAC's, but this issue is too hard to deal with. Here I assume that the discrepancy between AAC's and wholesale prices is small and can be ignored. Thus I simply treat AAC's as measuring wholesale prices.

I examine the retail and wholesale price patterns of StarKist and CKN in different stores. I find that wholesale prices display little variation across stores, but retail prices vary substantially across stores in different zones. Dominick's adopted zone pricing during the sample period (Chintagunta, Dube and Singh, 2002). I also find that a wholesale price reduction is always accompanied by a retail price reduction. However, retail price reductions frequently occur during periods when there is no change in wholesale prices. Therefore, it seems that most price promotions are retailer-driven. Figures 1.9 and 1.10 illustrate the retail and wholesale prices of StarKist and CKN over time in two Dominick's stores.

Based on the setup of supply side in the previous section, retailers incur losses of category profits from their sales. I then conduct a counterfactual experiment to quantify the potential losses due to retailer's price promotions.

First, I assume that the product category consists of only two brands, StarKist and CKN, and that their wholesale prices are the same and constant. I vary wholesale prices in the range from \$0.49 to \$0.64, which is about the range of regular wholesale prices observed in the DFF database. Second, given the wholesale prices, I compute the category-profit-maximizing retail prices (the "optimal" prices) and the demand using the demand-side estimates. Finally, I compare the retailer's category profits under two scenarios. In "Scenario 1", the retailer adopts the optimal prices for two weeks. In "Scenario 2", the retailer runs a price promotion for StarKist in one week and a price promotion for CKN in the other week where the non-sale price is \$0.89 and the sale price is \$0.69.

Table 1.11 reports the results of this counterfactual experiment. The last column of Table 1.11 shows that the loss of category profits increases in wholesale prices, as expected. The more interesting point in this counterfactual is that if the modal non-sale price (\$0.89) observed in the data is actually the optimal retail price (which implies the wholesale price is about \$0.54), then by running staggered sales on StarKist and CKN, the retailer incurs a loss of about 17% of the category profits.

I conduct another counterfactual experiment using Dominick's data to estimate the loss of category profits from retailers' cyclical pricing. This counterfactual is conducted as follows. First, for a given store, I sample 3000 households using the demand-side estimates and the distribution of household income in the zip code area where the store is located (I assume that the income distribution is lognormal and a store-level income distribution is estimated using the Census 2000 Data from the U.S. Census Bureau). Second, given the observed weekly retail prices, I simulate demand for StarKist and CKN by aggregating the demand of the 3000 sample households and then calculate the weekly category profits, i.e., "weekly data profits", using the observed weekly wholesale prices. Third, I compute the "total data profits" by summing the "weekly data profits" over the sample period. Fourth, given the observed weekly wholesale prices and the demand condition generated by the 3000 sample households, I compute the "weekly optimal prices" which maximize each store's weekly category profits. Given the "weekly optimal prices", I first simulate demand, and then calculate the weekly category profits, i.e., "weekly optimal profits", using the observed weekly wholesale prices. Finally, I compute the "total optimal profits" by summing the "weekly optimal profits" over the sample period. The ratios of "total data profits" over "total optimal profits" for 16 Dominick's stores are reported in the last column of Table 1.12.

The loss of category profits from retailers' cyclical pricing ranges from 19% to 44%, depending on a store's retail price pattern. In general, the loss is greater for the stores which have lower non-sale prices and hold less frequent sales (referred to as "every-day-low-price" stores) and the loss is smaller for the stores which have higher non-sale prices and hold more frequent sales (referred to as "high-low" stores). Figures 1.9 and Figure 1.10 illustrate the "optimal" price patterns in a "every-day-low-price" store and a "high-low" store respectively, and show that there is a tendency for my demand model to over-predict retail markups. Overestimation of markups and the heterogeneity in price patterns across stores suggest that retail price levels and price promotions are likely related to competition among retailers and its interaction with local market conditions.

1.9 Conclusion

Using scanner data on household purchases of canned tuna in Springfield, MO, I empirically examine the impact of price promotions on consumer demand. The finding is that market-expansion effects usually dominate the brand-switching effect. I then apply the demand estimates to the supply side to investigate the profitability of retailer price promotions. A counterfactual experiment on Dominick's data shows that the loss from retailers' cyclical pricing is not small. Judging from the tendency of my model to over-predict retail markups, and from the variation in markups across stores, retailer price promotions reflect local competition among stores for customers who find shopping around costly. Intertemporal demand effects such as inventory effects, state-dependent preferences and brand loyalty effects are ruled out as possible drivers of sales. Future investigations of the motivations fr retailer sales may therefore need to incorporate competition into the current model.

Chapter 2

The Effects of Store-Brand Entry on Consumer Demand and Welfare: An Empirical Study in the Hot-Breakfast Cereals Product Category

2.1 Introduction

One noticeable phenomenon in the supermarket industry in the recent two decades is the strong prevalence of store brand (SB) products in retailers' chain stores. According to the Private Label Manufacturers Association, in 2003, the total market share of SB's in supermarkets came in at 20.7% in terms of units and 16.3% in terms of dollars, and total sales reached $42.9 billion. One reason that retailers are spending resources on creating and maintaining their own brands is because a successful SB program is viewed as an advantageous competitive strategy to them. For example, since consumers can buy NB products everywhere, but they can only buy a certain retailer's SB products in its chain stores, SB's can help retailers differentiate themselves from their competitors. Some empirical analyses in the literature show that SB's contribute to greater store loyalty and increase retailers' profits (Corstjens and Lal, 2000; Sudhir and Talukdar, 2004). Moreover, if a considerable amount of current NB buyers switch to SB's after they enter the market, retailers may be able to increase their bargaining power against NB manufacturers. There is empirical evidence showing that retailers' bargaining position is strengthened after they introduce SB's (Narashimhan and Wilcox, 1998; Morton and Zettelmeyer, 2000). In this chapter, I study the effects of SB entry on consumer demand and welfare.

Both retailers and NB manufacturers may be interested in knowing how demand could be affected by the entry of SB's. The former have direct control over SB's and the latter face additional competition from SB's. In most cases, SB's are "me-too" products which imitate the leading NB's in a product category. Moreover, they are usually priced quite lower than NB's and have higher retail margins.[1] For retailers, by introducing the low-priced SB's into a product category, not only may it induce some current NB buyers to switch to SB's, but the total category sales may expand if SB's can attract new consumers who find NB's too expensive. However, even though the margins of SB's are generally higher than those of NB's, since their prices are lower, it is not necessarily the case that retailers can earn higher profits simply by stealing more business from NB's. Retailers need to understand the way SB entry affects the demand elasticities of incumbent NB's so that their pricing among NB's and SB's achieves the goal of profit maximization. As for NB managers, an understanding of whether consumers have shifted away from their brands to SB's and of the potential changes in consumers' price sensitivities given their current marketing activities could help them decide how to respond to SB entry, e.g., whether it is necessary to increase advertising expenditures to improve brand perception and strengthen brand loyalty in a given market area (Chintagunta et al, 2002).

The first step of my empirical approach to studying the effects of SB

[1] retail margin $= (\frac{\text{retail price} - \text{wholesale price}}{\text{retail price}})\%$

entry on consumer demand and welfare is to estimate consumer demand for NB's and SB's in a particular product category - hot-breakfast cereals - using scanner data from a multi-store supermarket chain in a major metropolitan area. The demand side is modeled by a random-coefficients logit model and estimated by the GMM approach where retail prices are instrumented to account for the potential problem of endogeneity. Since SB's are introduced during the sample period, given the demand-side estimates, I examine how the entry of SB's affects the demand elasticities of incumbent NB's by comparing the own-price elasticities of NB's before and after SB entry. I further study consumers' patterns of substitution between national and store brands (e.g., whether consumers tend to substitute within NB's/SB's, or substitute between SB's and their corresponding imitated NB's) by examining the cross-price elasticities of national and store brands. Finally, I quantify the changes in consumer welfare due to the entry of SB's. The empirical findings are: (1) After the entry of SB's, demand becomes more elastic for non-imitated NB's, and either more elastic or shows no change for imitated NB's; (2) in general, substitution patterns for NB's and SB's are asymmetric, i.e. when the prices of their favorite products increase, most NB buyers tend to substitute to other NB products, but SB buyers will substitute to the corresponding imitated NB's; (3) the increase in consumer surplus due to the entry of SB's is trivial for an individual consumer ($0.001/week assuming that consumers have a unit demand of one serving per week), but the aggregate benefit could be quite substantial ($5,566/week for the market under study).

The rest of this chapter is organized as follows. Section 2.2 provides an overview of the demand model. Section 2.3 describes the data. Section 2.4 discusses the estimation procedure. Section 2.5 presents the estimation results. Section 2.6 concludes.

2.2 Demand Model

The demand side is specified by a random-coefficients logit model which has two advantages over other types of demand models. First, suppose demand for a product category containing a large number of differentiated products, n, is to be estimated. If no restriction is imposed, we'll need to estimate at least n^2 parameters in a log-log demand model. The logit specification can reduce the dimensionality problem if consumer preferences can be projected onto a set of exogenous product characteristics. Moreover, the additional random-coefficients specification allows more flexible own-price and cross-price elasticities. In a standard logit model, products with the same market share have the same own-price demand derivative, which implies the same markup in a single-product-firm setting. However, market share should not be the only factor that determines a product's markup. Instead, in a random-coefficients logit model, own-price demand derivatives are more flexible because they are driven by the price sensitivities and purchase probabilities of heterogeneous consumers. In other words, the independence-of-irrelevant-alternatives (I.I.A.) property in a logit model imposes strong restrictions on the substitution patterns among the products. By adding a random component to the coefficients, cross-price elasticities are driven by the products' characteristics and consumers' preferences over those characteristics, which generates more realistic substitution patterns (Berry, Levinsohn, and Pakes, 1995; Nevo, 2000).

The demand side is specified as follows. Suppose there are T markets, and in each market, there are I_t consumers and J_t products. Let t denote market, $t = 1, \ldots, T$; i denote consumer, $i = 1, \ldots, I_t$; and j denote product,

$j = 1, \ldots, J_t$. Here a market is defined as a store-week combination. Consumer i's conditional indirect utility from product j at market t is

$$u_{ijt} = x_j\beta_i + \alpha_i p_{jt} + \lambda SD_t + \gamma d_{jt} + \xi_j + \triangle\xi_{jt} + \varepsilon_{ijt} \tag{2.1}$$

Here x_j is a vector of product j's observed product characteristics and β_i is consumer i's taste for x_j; p_{jt} is the price of product j at market t and α_i is consumer i's price sensitivity or marginal utility of income; SD is a dummy variable for the summer season (June, July and August) which controls for the possible seasonal effects on utility for the product category and λ is the parameter on SD; d_{jt} is a dummy variable indicating whether product j is on promotion at market t and γ measures consumers' sensitivity to promotions; λ and γ are assumed be the same for all consumers. The parameter ξ_j is the mean valuation of product j's unobserved product characteristics and $\triangle\xi_{jt}$ is a market specific deviation from this mean; $\triangle\xi_{jt}$ may be caused by factors such as shelf space, shelf location, or other unobserved demand drivers which vary across stores and weeks; finally, ε_{ijt} is the idiosyncratic taste which is independently and identically distributed across consumers, products, and markets and follows a Type I extreme-value distribution.

The individual-specific parameters are assumed to follow a multivariate normal distribution conditional on consumers' demographics, i.e.,

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma\nu_i, \tag{2.2}$$

$$\text{where } \nu_i \sim N(0, I)$$

Here D_i is a vector of observed consumer characteristics of consumer i and Π is a matrix of parameters measuring how consumers' preferences vary with demographics; ν_i is a vector drawn from a multivariate standard normal distribution and is interpreted as unobserved consumer characteristics affecting consumers' preferences through a scaling vector Σ. Thus, the individual-specific parameters $\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$ follow a normal distribution with mean $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i$ and a diagonal covariance matrix Σ^2.

In addition to the J_t "inside" goods at market t, an "outside" good has to be introduced in order to complete the demand system. A choice to purchase the outside good means that a consumer chooses not to purchase any of the J_t inside goods. According to Berry (1994), without the outside good option, a homogeneous increase in prices of all inside goods will not cause a decrease in total category demand, i.e., category demand is perfectly inelastic. This is unreasonable since product categories can be substitutes for each other (if the definition for a product category is not too broad). Therefore, demand for a product category is usually not perfectly inelastic. For identification purposes, the indirect utility from the outside good is normalized to $u_{i0t} = 0 + \varepsilon_{i0t}$.

Following Nevo (2001), combining Equation (2.1) and Equation(2.2) results in:

$$u_{ijt} = \delta_{jt}(x_j, p_{jt}, \xi_j, \triangle\xi_{jt}; \theta_1) + \mu_{ijt}(x_j, p_{jt}, \nu_i, D_i; \theta_2) + \varepsilon_{ijt} \tag{2.3}$$

Where

$$\delta_{jt} = x_j\beta + \alpha p_{jt} + \lambda SD_t + \gamma d_{jt} + \xi_j + \triangle\xi_{jt}$$

$$\mu_{ijt} = \begin{pmatrix} p_{it} & x_j \end{pmatrix} (\Pi D_i + \Sigma\nu_i)$$

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \theta_1 = \begin{pmatrix} \alpha & \beta' & \lambda & \gamma \end{pmatrix}', \theta_2 = \begin{pmatrix} vec(\Pi)' & \Sigma' \end{pmatrix}'$$

The indirect utility in Equation (2.3) consists of two parts: δ_{jt} and $\mu_{ijt}+\varepsilon_{ijt}$. δ_{jt} represents the mean utility, which does not vary across consumers at market t and $\mu_{ijt} + \varepsilon_{ijt}$ is a heteroskedastic deviation from the mean utility, which captures the effect of random coefficients. The vector θ denotes all of the parameters to be estimated in the model.

Assuming that consumers have unit demand and choose a product which maximizes utility and that ties occur with zero probability, product j's share in market t is

$$s_{jt}(x, p_t, \delta_t; \theta_2, F^*) = \int_{A_{jt}} dF^*(D, \nu, \varepsilon) \tag{2.4}$$

Where

$$A_{jt}(x, p_t, \delta_t; \theta_2) = \{(D_i, \nu_i, \varepsilon_{it})|\ u_{ijt} \geqslant u_{ilt}, \text{ for } l = 0, 1, \ldots, J_t\}$$

Here x, p_t, and δ_t are, respectively, observed product characteristics, prices and mean utility levels of the inside goods at market t, and $F^*(D, \nu, \varepsilon)$ is the joint distribution of individual attributes.

Given the values of the parameters and the joint distribution of individual attributes, Equation (2.4) can be computed either analytically or numerically. One approach to estimation is to choose the parameter values which minimize the distance between the predicted market shares derived from Equation (2.4) and the observed shares in the data. To do so, two problems have to be addressed first. The first comes from the fact that market size is usually not directly observed. According to Berry (1994), most of the time, researchers only have information on quantities being purchased for the J_t inside goods at each market. In order to compute the $J_t + 1$ market shares, i.e., market shares for both inside and outside goods, total market size must be estimated. Nevo (2000) makes market size proportional to an observed measure of market population. He suggests that this measure should be chosen so that the potential market is large enough to allow for a non-zero outside good market share and so that the estimated results should not be too sensitive to the definition. The second problem is that prices are potentially correlated with the unobserved demand shocks, $\xi_j+\Delta\xi_{jt}$. Therefore, in order to consistently estimate the parameters, prices have to be instrumented by variables which are correlated with them but uncorrelated with the unobserved demand shocks. Regarding the first problem, I define the potential market size as a store's average number of visiting customers each week multiplied by the average household size of the zip code area where the store is located. As for the second problem, I use lagged retailer average acquisition costs as the instrumental variables. Details are provided in sections 2.4.1 and 2.4.2.

2.3 Data Description

The two major databases used for estimation are the Dominick's Finer Foods data from the James M. Kilts Center, GSB, University of Chicago and Census 2000 Data from the U.S. Census Bureau.

2.3.1 Dominick's Finer Foods Data

Dominick's Finer Foods (DFF) is the second largest supermarket chain in the metropolitan area of Chicago, Illinois. DFF has about 100 stores and has a market share of around 25% in this area (Chevaleir, Kashyap and Rossi, 2000). The variables in the DFF data include retail prices, unit sales, retail profits, and a deal code indicating whether a product was sold on a promotion. The data are provided on a weekly basis from September 1989 to May 1997 for each UPC in 29 different product categories and for each store. Weekly store traffic is also available and is defined as the total number of customers who visited the store and made a purchase during a week. One feature of the DFF data is that it reveals information on the retailer's profit margins through the retail profits variable which is defined as the percentage gross margins DFF makes on the dollar sales for each UPC. Therefore, retail costs can be recovered by multiplying retail prices with retail profits. However, according to the data provider, the costs recovered this way are average acquisition costs (AAC) of the items in inventory and they do not exactly correspond to the retailer's marginal (or replacement) costs.

The particular product category studied here is hot-breakfast cereals. There are data on 96 UPC's from June 1991 to April 1997 (304 weeks) for each store in this product category. However, the data are not balanced, i.e., not all the UPC's are offered in all the stores all the time. Entries of new UPC's and exits of old ones are observed during the sample period. To reduce the number of inside goods, I first select 23 of the 96 UPC's and then aggregate some of them together, which generates 11 "final products" to be included in the demand system. The following describes how the 23 UPC's are chosen. I first select all SB's in this product category (6 UPC's). Then for NB's, I only select those which are available in the market during the whole sample period and have relatively high category shares (17 UPC's). The reason that I aggregate over some subsets of the 23 UPC's is because the UPC's in the same subset not only share very similar product characteristics, but their prices are almost always the same. Therefore, aggregation does not seem to hurt the identification of the key parameters we are interested in and makes the demand model more parsimonious.

Towards the end of the sample period, the retailer introduced a total of 6 SB's into this product category. 5 SB's entered the market in the 123rd data week and the sixth SB in the 203$^{\text{rd}}$ data week. Since I am interested in studying the effects of SB entry on consumer demand, I include all SB's in the demand system. But because the SB introduced last is to be aggregated with two of the SB's introduced earlier, I didn't use the data from the 80-week period between the introduction of the first five SB's and the sixth SB. In my model, I assume that consumers have perfect information on all products in the same category. Therefore, I would like to treat this period as some sort of learning period (i.e., since it is the first time SB products were introduced into this category, consumers need some time to try and learn about their qualities) to justify the decision of disregarding the data during the learning period. In addition, I assume that consumers know the quality of the SB introduced last from their previous usage of the two SB's which are in the same product line as the 6$^{\text{th}}$ SB. Thus no more learning time is needed. Therefore, the data used for estimation can be separated into two periods: the first period (the "before" SB entry period) is before any of the 6 SB's entered the market, which is from data week 4 to 122 (06/27/91 - 10/13/93), and the second period (the "after" SB entry period) is after all 6 SB's entered, which is from data week 206 to 304 (05/25/95 - 04/30/97).

Table 2.1 contains basic information about the 23 UPC's included in the demand system, including UPC number, product name, unit weight (measured by oz), serving size, and the "product" into which the UPC's are to be aggregated. I use serving as the measurement unit and convert sales volume into number of servings to compute each product's market share. The serving size per container is based on the table of nutrition facts printed on the package. One reason for using serving, instead of weight, as the measurement unit is that the UPC's that are aggregated together have the same serving size even though their unit weights are different. For instant hot-breakfast cereals, the serving sizes are intuitive: the serving size per container corresponds to the number of individually wrapped packs per container. As for non-instant ones, the packages are one single box or tube without separately wrapped small packs. Therefore, I assume that consumers' consumption amount follows the per serving size suggested by the manufacturers.

After aggregation, there are 11 products in the demand system. These products can be differentiated in several ways, including type of grain (wheat or oats), ease of preparation (instant or non-instant), packaging (small packs, regular box/tube, or big tube), flavor (original or flavored), and manufacturers/brands (Nabisco, Quaker, or Dominick's). Table 2.2 lists the 11 products and their observed product characteristics. Products 1 to 7 are NB's (i.e., with brand name either Nabisco or Quaker) and Products 8 to 11 are SB's (i.e., with brand name Dominick's). The coefficient on the product characteristic, "constant", can be interpreted as consumers' preferences for Quaker's non-instant, original oats with regular packaging (referred as the "base product") relative to the outside good. Because all Nabisco-manufactured products are wheat and all wheat products are manufactured by Nabisco, consumers' preferences for wheat and Nabisco cannot be identified separately, i.e., only the coefficient on "wheat/Nabisco" can be identified. Similarly, because all the instant products have the packaging of individually wrapped packs and vice versa, only the coefficient on "instant/small-pack" can be identified. Finally, Products 8, 9, 10 and 11 are the SB versions of the NB products, Products 4, 5, 6, and 7; each pair has the same product characteristics except the characteristic, "SB". Therefore, the coefficient on "SB" can be viewed as consumers' valuation of the brand name - Dominick's relative to the brand name - Quaker, which may be affected by factors such as vertical quality differences between SB's and imitated NB's or the strength of brand image (due to advertising etc.).

2.3.2 Census 2000 Data

In the random-coefficients specification, I allow individual parameters to vary with demographics. In practice, I use two observed consumer characteristics to interact with the observed product characteristics and prices, i.e., logarithm of household income ("logincome") and a dummy variable ("hhund18") which equals one if a household has members under eighteen and zero otherwise. The distributions of the two demographic variables are estimated using the Census 2000 Data from the U.S. Census Bureau. I select sixteen of the 96 stores owned by DFF, which are in operation during the whole sample period and have relatively fewer missing data. Information about each store's location, including address, zip code, and city is available in the data. Since Census 2000 Data are accessible at the zip code level, I use them to create store-level demographic distributions. The distribution of household income in the census data is discrete, given as the percentage of households in ten income intervals. To ease the process of sampling household income when simulating market shares, I assume that the distribution of household income is log-normal. I then estimate the parameters (mean and variance) by the least squares method, i.e., by fitting the estimated log-normal distribution to the empirical distribution in the Census data. As for the distribution of "hhund18", I assume that it follows a Bernoulli distribution. I simply take the sample proportion of households with individuals under age 18 in the Census data as the parameter. Moreover, because I don't have information on the joint distribution of "logincome" and "hhund18", I assume that the marginal distributions on these two variables are independent. For each store, 63 samples are drawn from the store-level demographic distributions by Halton sequences, and the same sets of samples are used for these stores over time. Table 2.3 lists the 16 stores and the store-level demographics.

2.4 Estimation Procedure

2.4.1 Market Shares

Serving is used as the measurement unit to define market shares. In order to construct market shares for both inside and outside goods, market size has to be defined. The DFF data contain information on weekly store traffic defined as the number of customers who visit the store and make a purchase during a week. I assume that each household visits a store once a week and that demand for this product category is one serving per household member per week. For each store, I first compute the average weekly store traffic over the sample period. Then I define the potential market size for a store as its average weekly traffic multiplied by the average household size of the zip code area where the store is located. The last two columns in Table 2.3 report the average household size and store traffic of the sixteen stores. I use average weekly store traffic, instead of weekly store traffic, to compute market size since it is more likely that weekly store traffic can be correlated with prices. It seems more reasonable to assume that average weekly store traffic and thus market size is exogenous with respect to prices of inside goods. By defining market size in this way, market size is time invariant for each store. Given the potential market size, I convert the unit sales of each product into total number of servings and divide them by the potential market size to derive the market shares for the inside goods. Since market shares should sum to one, the market share of the outside good is defined as the difference between one and the sum of the market shares of the inside good. Finally, besides sales volume, prices and costs are converted into prices and costs per serving as well and then deflated by the yearly CPI in the Chicago-Gary-Kenosha area (the CPI data are from the U.S. Department of Labor). In the case when a product is the aggregation result of more than one UPC, I use volume-weighted average price.

2.4.2 Instruments

Because prices are potentially correlated with unobserved demand shocks, instrumental variables for prices are needed in order to get a consistent estimator. Cost variables are traditionally appropriate instruments (Berry, 1994). Moreover, lagged values of these variables also provide natural candidates in many time-series settings (Greene, pp 375). Therefore, I use lagged average acquisition costs (AAC's) as the instruments for prices. AAC is the average inventory cost of an item held by a store each week and is recovered by multiplying retail prices with retail profits. Obviously, AAC's and prices are correlated. Therefore, whether they are valid IV's depends on whether they are correlated with unobserved demand shocks. One possibility that causes AAC's to be correlated with unobserved demand shocks is when the observed product characteristics specified by researchers are not thorough enough to capture all the factors which determine utility. Then utility derived from the "unobserved" product characteristics will become part of unobserved demand shocks. If these product characteristics are only "unobserved" to researchers, not manufacturers, manufacturers will likely take them into account when pricing their products, which implies that wholesale prices, as well as AAC's, will be correlated with unobserved demand shocks. The solution to this problem is to include product-specific dummy variables in the product characteristics space. According to Nevo (2001), the coefficient on a dummy variable for, say, product j captures the part of the mean utility which does not vary with markets, i.e., $x_j\beta + \xi_j$, so only the market-specific deviation, i.e., $\triangle\xi_{jt}$ is left in the unobserved demand shock for product j at market t. However, if there is serial correlation in $\triangle\xi_{jt}$'s, since AAC's depend on current and lagged wholesale prices and lagged sales, they will be correlated with unobserved demand shocks (Chintagunta, 2002). Therefore, I assume that market-specific deviations are not serially correlated.

2.4.3 Computation

Following Nevo's computational algorithm (Nevo 2000 and Nevo 2001), I construct a GMM estimator for demand-side parameters. The GMM estimate is

$$\hat{\theta} = \arg\min_{\theta} \; \mathrm{e}(\theta)' Z W Z' e(\theta),$$

$$\text{where } \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \; \theta_1 = \begin{pmatrix} \alpha & \beta' & \lambda & \gamma \end{pmatrix}', \; \theta_2 = \begin{pmatrix} vec(\Pi)' & \Sigma' \end{pmatrix}'$$

Here Z is a set of instruments exogenous to the unobserved demand shocks; $e(\theta)$ is an error term in the regression model and is a function of the parameters; and W is a weighting matrix. The parameters in θ_2 are the interaction terms between D_i/ν_i (i.e., observed/unobserved consumer characteristics) and X_2 (i.e., price and observed product characteristics, including constant, wheat/Nabisco, instant/small-pack, big tube, flavor, and SB); and θ_1 contains the parameters on X_1 (i.e., price, summer dummy, deal dummy, and product-specific dummies). Note that θ_2 enters the GMM objective function nonlinearly and θ_1 linearly.

The computational algorithm consists of 4 steps.

Step 1: Given θ_2, compute the mean utility levels, δ_t's, market by market, which solve the implicit system of equations described by Equation (2.5).

$$s_t(X_2, \delta_t; \theta_2) = s_t^0, \qquad \text{for } t = 1, \ldots, T \tag{2.5}$$

Here $s_t(.)$ is the market-share function defined by Equation (2.4) and s_t^0 are the observed market shares in the data. Since Equation (2.4) does not have a closed-form solution, a simulation estimator is used to approximate the predicted market shares. First, pairs of (ν_i, D_i), for $i = 1, \ldots, ns$, are drawn from F_{ns} (i.e., a multinomial normal distribution and the distributions on demographics), which represent unobserved and observed consumer characteristics. The simulated market shares are then computed according to Equation (2.6) and Equation (2.7).

$$s_{jt}(X_2, \delta_t, \theta_2, F_{ns}) = \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \tag{2.6}$$

$$s_{ijt}(X_2, \delta_t, \theta_2, F_{ns}) = \frac{\exp[\delta_{jt} + X'_{2j}(\Pi D_i + \Sigma \nu_i)]}{1 + \sum_{l=1}^{J_t} \exp[\delta_{lt} + X'_{2l}(\Pi D_i + \Sigma \nu_i)]}, \tag{2.7}$$
$$\text{for } i = 1, \ldots, ns,\ j = 1, \ldots, J_t,\ t = 1, \ldots, T$$

In BLP (1995), they show that under regular conditions, for any triple of (s_t^0, θ_2, F), δ_t in Equation (2.5) can be solved recursively by iterating a contraction mapping operator until convergence is reached. A series of $\{\delta_t^n\}_{n=1}^N$ is first computed numerically according to the relation: $\delta_t^{n+1} = \delta_t^n + \ln(s_t^0) + \ln(s_t(X_2, \delta_t^n, \theta_2, F_{ns}))$. Then δ_t is approximated by δ_t^N, where N is the smallest integer such that $\|\delta_t^N - \delta_t^{N-1}\|$ is smaller than some defined tolerance level.

Step 2: Given θ_2 and $\delta(\theta_2) = \begin{pmatrix} \delta'_1 & \ldots & \delta'_T \end{pmatrix}'$, an error term is defined as $e(\theta) = \delta(\theta_2) - X_1\theta_1$. Then interact the error term with the instruments and the weighting matrix to construct the GMM objective function, i.e.,

$$q = e(\theta)' Z W Z' e(\theta) \tag{2.8}$$

Step 3: Search for θ_2 to minimize Equation (2.8). Note that since $e(\theta)$ is linear in θ_1, given θ_2 and δ, the value of θ_1 which minimizes q can be solved analytically. The solution is a function of θ_2, i.e.,

$$\theta_1(\theta_2) = (X_1' Z W Z' X_1)^{-1} X_1' Z W Z' \delta(\theta_2) \tag{2.9}$$

Therefore, searching for the demand-side parameters can be reduced to the search for θ_2 only.

Step 4: Recover the taste coefficients, β, by a GLS regression, i.e., by regressing $\hat{\rho}$, the estimated coefficients on the product-specific dummies, on X, the observed product characteristics, using a weighting matrix V_ρ^{-1}, which is inverse of the variance-covariance matrix of $\hat{\rho}$. The GLS regression model is

$$\rho = X\beta + \xi \tag{2.10}$$

Assuming $E[\xi|X] = 0$, the GLS estimator for β is

$$\hat{\beta} = (X' V_\rho^{-1} X)^{-1} X' V_\rho^{-1} \hat{\rho} \tag{2.11}$$

2.5 Estimation Results

2.5.1 Descriptive statistics of the data

Table 2.4 reports the average quantities, retail prices, AAC's, and retail profits for each product over the before and the after SB entry sample periods. Regarding quantities, after the entry of SB's, except for Product 2, the quantities of all other NB's decreased. Note that the increased sales of SB's are less than the decreased sales of NB's. Therefore, it seems that introducing SB's only induced consumers who used to buy the incumbent NB's in this product category to switch to the SB's, but did not expand the total category sales by attracting new consumers into the market. Regarding retail prices, after the entry of SB's, retail prices of the 4 imitated NB's all decreased and those of the 3 non-imitated NB's all increased. AAC's displayed the same pattern, except for Product 1. If the retailer's marginal costs consist of only wholesale prices and if wholesale prices can be approximated by AAC's, then the changes in retail prices could be driven by the changes in wholesale prices. A possible explanation for the cost-side adjustments may be that since SB's compete directly and thus steal most of business from their imitated NB's (which are the leading brands in this product category), the manufacturer of these imitated NB's (Quaker) may want to fight for market share by dropping wholesale prices in the hope that the retailer will pass on the savings to consumers by decreasing retail prices. However, it seems odd that the wholesale prices of two of the non-imitated NB's increased after SB entry and that of the third one decreased just a bit. My guess is that even though the non-imitated NB's also lost business to SB's, since they are not major players in this product category (the ratio of sales of the non-imitated NB's to sales of the imitated NB's before SB entry is about 1 to 6), the loss may be quite limited. Moreover, the decrease in retail prices of the imitated NB's had made the competition even more severe. The manufacturer of the non-imitated NB's (Nabisco) may find that competing via price is not profitable. Therefore, instead, it adopts a strategy of further differentiating their brands from the low-priced SB's by positioning them as the premium brands and maintaining a higher-price level. Finally, no matter what the manufacturers' pricing strategies are after SB entry, introducing SB's into the market seems to have a positive impact on the retailer's profitability since retail profits of NB's all increased after the entry of SB's.

The bottom panel in Table 2.4 compares SB's and their corresponding imitated NB's. The average retail prices of SB's are about 76%-81% of those of NB's and the average sales are 18%-41%, which implies that even though SB's are less expensive and are "me-too" products to the imitated NB's, some sort of perceived quality differences do exist so that a large number of consumers are willing to pay more for the imitated NB's. Therefore, even though the retail profits of SB's are much higher than those of NB's, which seems to suggest that the retailer should raise the prices of NB's to steal more business and thus make more profits, the retailer need to consider the possibility that by doing so, some consumers who are loyal to NB's or who find the quality of SB's unacceptable may decide to buy nothing in this category, or worse, switch to other retailers. Hence, it is important for retailers to know consumers' demand elasticities so that their pricing decisions indeed achieve the goal of profit maximization.

2.5.2 Results of the logit model

Table 2.5 presents the estimation results for the logit model. In the standard logit specification, consumers are homogeneous, i.e., their taste preferences are the same. Therefore, there are no random coefficients. This makes the estimation very simple since the mean utility level, δ_{jt}, can be solved analytically, i.e., $\ln(s_{jt}^0) - \ln(s_{0t}^0)$. Columns (i) and (ii) shows the results of regressing $\ln(s_{jt}^0) - \ln(s_{0t}^0)$ on retail price, summer, promotion, and the observed product characteristics (Set A covariates) by OLS and 2SLS respectively. Columns (iii) and (iv) shows the results of regressing $\ln(s_{jt}^0) - \ln(s_{0t}^0)$ on retail price, summer, promotion, and product-specific dummies (Set B covariates) by OLS and 2SLS respectively.

Most of the estimated coefficients are of the expected sign. For example, the negative coefficient on the seasonal dummy, "summer", makes sense since the product category is *hot*-breakfast cereals. The coefficient on "SB" is negative since SB products are usually perceived as of lower quality. The negative coefficient on "big" implies consumers dislike products in large-size packages and it is probably because they are harder to store or transport (Cohen 2004). The negative coefficients on the constant and the product-specific dummies may reflect the fact that the market shares of the inside goods are very small relative to the market share of the outside good. Except for the coefficient on instant/small-pack in the OLS regression, the coefficients on instant/small-pack and on flavor are positive. The positive sign may be caused by consumers' treating these two product characteristics as some sort of product innovation or improvement which enhances utility. The coefficient on wheat/Nabisco is negative, which is bad news for the manufacturer of Nabisco because it implies that consumers do not regard Nabisco as a premium brand in this product category. Finally, the coefficient on price is negative as expected.

Note that when Set A covariates are used as the regressors, the price coefficient (in absolute value) estimated by OLS (denote as α^{OLS}) is less than that estimated by 2SLS (denote as α^{2SLS}), but when Set B covariates are used, the opposite occurs. When the regressors are Set A covariates, the error term is $\xi_j + \triangle\xi_{jt}$, but when regressors are Set B covariates, the error term is $\triangle\xi_{jt}$. Therefore, α^{OLS} being greater than α^{2SLS} using Set B covariates seems to imply a negative correlation between $\triangle\xi_{jt}$ and prices, which causes the upward bias of α^{OLS} when the endogeneity of price is not accounted for. I am not sure how to explain the negative correlation between $\triangle\xi_{jt}$ and prices. On the other hand, a positive correlation between ξ_j and prices seems more intuitive. α^{OLS} being less than α^{2SLS} using Set A covariates may be due to the positive correlation between ξ_j and prices which dominates the negative correlation between $\triangle\xi_{jt}$ and prices, so that α^{OLS} is downward biased when the endogeneity of price is not accounted for.

Judging by the adjusted R^2's of OLS, it seems that including the product-specific dummies improves the fit of the model. The J statistics of 2SLS look more worrisome. The over-identification test is rejected using either Set A or B covariates, which suggests that the identification assumption, i.e., that the instruments are uncorrelated with the error terms, is not valid. However, Nevo (2001) comments that a χ^2 test will reject essentially any model with a large enough sample size. Since I have about 27,700 observations, which should be counted as a large sample size, the failure of the over-identification test may not be conclusive enough in rejecting the identification assumption.

In the logit model, product j's own-price elasticity in market t is $\alpha(1 - s_{jt})p_{jt}$. I use α^{2SLS} with Set B covariates to compute the own-price elasticities. Table 2.6 reports the medians of own-price elasticities for the before and after SB entry periods. After the entry of SB's, demand for both imitated and non-imitated NB's becomes more elastic and the percentage changes in own-price elasticities are greater for the non-imitated NB's than for the imitated ones. Since the entry of SB's provides consumers with more alternative options in this product category, it seems reasonable to expect more elastic demand.

Even though the logit model has the advantage of being computationally simple, the I.I.A. property results in restrictive and unrealistic substitution patterns. For example, when there is a change in price of, say, product

j, the proportions of consumers who substitute away from product j to other products are only based on the market shares of these products. Therefore, cross-price elasticities in the logit model are not reported here. In fact, as suggested by Nevo (2001), it seems that the own-price elasticity estimates of a logit model can also be quite restrictive, since some of the products in Table 2.6 have own-price elasticities in the inelastic region.

2.5.3 Results of the random coefficients logit model

The random coefficients logit model (the full model) is described by Equation (2.3) and the model parameters are estimated according to the procedure in Section 4.3. Table 2.7 reports the estimation results for the full model. The results in the first column are consumers' mean tastes for price (α), observed product characteristics (β's), summer (λ), and promotion (γ). The results in the second, third, and fourth columns are the coefficients on the interaction terms between price/observed product characteristics and unobserved/observed individual characteristics (Σ and Π).

The signs of the mean tastes for product characteristics: "inst/small-pack", "big", "flavor", and "SB" seem to be counter intuitive based on the arguments in Section 5.2. However, since the logarithm of household income is never zero, the signs of the individual-specific tastes for these product characteristics may change after accounting for the random coefficient effects. The signs of the rest of the coefficients seem to make more sense. For example, the coefficients on summer and promotion are, respectively, negative and positive; price sensitivity decreases with income and in households with individuals under 18; the marginal valuation for "inst/small-pack", "flavor", and "constant" increases with income and in households with individuals under 18; marginal valuation for "big", "wheat", and "SB" decreases with income and with household with individuals under 18. One implication from the random-coefficients results is that SB's seems more likely to be favored by lower-income people.

Table 2.8 reports the median own-price elasticities and the 95% confidence intervals for the full model, computed as follows. (i) I take 3000 random draws from the estimated distribution of the parameters. (ii) Given the parameters from each draw, I compute the own-price elasticities for each product at each market. (iii) I take medians for the before and after SB entry periods and compute the percentage changes of the two medians for each product. (iv) I order the simulated own-price elasticities for each product and construct the lower and upper bounds of the 95% confidence interval by taking the values in the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles.

Based on the point estimates, the own-price elasticities of NB's all increased after the entry of SB's, but the increase is not significant for Products 4 and 5. In theory, SB entry can affect a NB's demand elasticity in two potential ways. On the one hand, since the entry of SB's increases the number of substitutes to the incumbents, demand should become more elastic (referred to as the "substitute" effect). On the other hand, since SB's have lower prices than NB's, if consumers who switch to the cheaper SB's are price sensitive consumers and those who keep purchasing the more expensive NB's are brand-loyal consumers, demand elasticities of the incumbent NB's may decrease after the entry of SB's (referred to as the "loyalty" effect). Based on the above arguments, the results in Table 2.8 suggest that the substitute effect weakly dominates the loyalty effect. Moreover, similar to the results in the logit model, demand elasticities of the non-imitated NB's not only increased after SB entry, they increased relatively more than those of the imitated NB's, except for Product 7. A possible explanation is that after the entry of SB's, the non-imitated NB's face not only more competition from the new entrants, but more severe price competition from the imitated NB's as well. Therefore, even though SB's are used to target mainly the market of the leading brands, the non-imitated NB's can be hurt even worse by SB entry, due to the way the imitated NB manufacturer responds to the SB's.

Table 2.9 presents the median own-price and cross-price elasticities before and after the entry of SB's in the full model. One way to examine consumers' substitution patterns among these products is to compare the estimates in Table 2.9 by rows. The non-imitated NB's, Products 1-3, are most sensitive to a change in the price of the imitated NB's with similar product characteristics, i.e., Products 4, 6 and 7 respectively. Three of the imitated NB's, Products 4-6, are most sensitive to a change in the price of the other imitated NB's, i.e., Products 7, 4 and 7 respectively. Only one of the imitated NB's, Product 7, is most sensitive to a change in the price of its imitating SB, Product 11. As for the SB's, Products 8-11, they are most sensitive to a change in the price of their corresponding imitated NB's, Products 4-7. The results demonstrate the leading role of the imitated NB's in this product category and suggest why the imitated NB manufacturer has a stronger incentive to price aggressively in response to SB's than the non-imitated NB manufacturer.

Table 2.10 presents the medians of absolute change in market share with respect to percentage change in price before and after the entry of SB's in the full model. Another way to examine consumers' substitution patterns in this product category is to compare the estimates in Table 2.10 by columns. Increasing by 1% the price of the non-imitated NB's, most consumers substitute to the imitated NB's with similar product characteristics. Increasing by

1% the price of the imitated NB's, most consumers substitute to the other imitated NB's, except for Product 7. Increasing by 1% the price of the SB's, most consumers substitute to their corresponding imitated NB's. Therefore, the substitution patters are very similar to the ones based on Table 2.9. In general, substitution patterns for NB's and SB's are asymmetric, i.e. when the prices of their favorite products increase, most NB buyers tend to substitute to other NB products, but SB buyers will substitute to the corresponding imitated NB's. However, Product 7 seems to be an exception. A possible explanation is that consumers may find that no other NB's are a closer substitute to Product 7 than Product 11, or in terms of product characteristics space, the "physical" product characteristic, "flavor", dominates the "brand image" product characteristic, "SB". Therefore, when the price of Product 7 is increased, consumers are more likely to substitute to Product 11, the me-too SB version of Product 7. This suggests that simply imitating a NB does not necessarily make the imitating SB the closest substitute to the imitated NB. There are other factors the retailers should take into account when they introduce the "me-too" SB's to target the leading NB's market.

Finally, I examine how well the full model overcomes the restrictions on the own-price and cross-price elasticities imposed by the logit model. First, as pointed out by Nevo (2001), the problem with the own-price elasticities in the standard logit model is that they are determined directly by the functional form. For example, if the market shares of the inside goods are all fairly small (as they are in my case), $\alpha(1-s_{jt})$ will be nearly constant. Then the ordering of own-price elasticities, i.e., $\alpha(1-s_{jt})p_{jt}$, will follow closely the ordering of prices (Nevo, 2000). Table 2.11 shows that the median own-price elasticities in the logit model have exactly the same order as the prices. As for the median own-price elasticities in the full model, even though in general, higher-price products tend to have higher own-price elasticities and vice versa, the ordering of own-price elasticities is no longer identical to the ordering of prices. Second, in the logit model, when there is a change in the price of, say, product j, the proportions of consumers who substitute from product j to the other products are based on the market shares of these other products only, not on how close they are to product j in the product space. This is the result of the I.I.A. property and it implies that the cross-price elasticities in each column of the own- and cross-price elasticity matrix will be the same. Note that the ratios of the maximum to the minimum values of cross-price elasticities in each column of Table 2.9 are at least 14. These variations suggest that the full model does allow for more flexible substitution patterns. Therefore, the random-coefficients logit model seems to do quite well at overcoming the restrictions on the own- and cross-price elasticities imposed by the logit model.

2.5.4 Marginal costs

Given demand-side estimates, marginal costs can be backed out under certain assumptions about retailers' pricing behavior. I assume that marginal costs are constant and exogenous and that the retailer chooses prices to maximize the category-level profits for each market. The profit-maximizing problem of the retailer is described as follows.

$$\max_{p_t} \Pi_t(p_t) = M_t \sum_{j=1}^{J_t} (p_{jt} - mc_{jt}) s_{jt}(p_t), \qquad \text{for } t = 1, \ldots, T$$

Here p_t is a vector of retail prices of all products in the category; p_{jt}, mc_{jt} and s_{jt} are the retail price, marginal cost and market share of product j at market

t; and M_t is the potential market size of market t.

The solution to the above retailer's problem must satisfy J_t first order conditions as described by Equation (2.12).

$$s_{jt}(p_t) + \sum_{k=1}^{J_t}(p_{kt} - mc_{kt})\frac{\partial s_{kt}}{\partial p_{jt}} = 0, \qquad \text{for } j = 1, \ldots, J_t \tag{2.12}$$

These first order conditions can be rewritten in a matrix form as Equation (2.13).

$$s_t(p_t) + \Omega_t(p_t - mc_t) = 0 \tag{2.13}$$

Where $s_t = \begin{pmatrix} s_{1t} \\ \vdots \\ s_{J_t t} \end{pmatrix}$, $\Omega_t = \begin{pmatrix} \frac{\partial s_{1t}}{\partial p_{1t}} & \cdots & \frac{\partial s_{J_t t}}{\partial p_{1t}} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_{1t}}{\partial p_{J_t t}} & \cdots & \frac{\partial s_{J_t t}}{\partial p_{J_t t}} \end{pmatrix}$, $p_t - mc_t = \begin{pmatrix} p_{1t} - mc_{1t} \\ \vdots \\ p_{Jt} - mc_{Jt} \end{pmatrix}$

Given demand-side estimates, marginal costs can be backed out using Equation (2.14).

$$mc_t = p_t + \Omega_t^{-1} s_t(p_t), \qquad \text{for } t = 1, \ldots, T \tag{2.14}$$

Table 2.12 reports the average marginal costs (MC's) backed-out from the full model and the average acquisition costs (AAC's) observed in the data for the before and after SB entry periods. There are several explanations for the discrepancy between MC's and AAC's. For example, a temporary wholesale price cut this period may only show in the AAC's gradually as the old higher-priced inventory carried by the retailer is being sold over the next few periods. In this case, MC's are less than AAC's. On the other hand, if the retailer forward-buys to build up inventory during a manufacturers' trade promotion, AAC's will remain depressed for some time even after wholesale prices have gone back up (Chevalier et al, 2000). In this case, MC's are greater than AAC's. In addition, the nature of AAC can cause the discrepancy between MC's and AAC's as well. AAC is the average wholesale price of the item in inventory. According to Chintagunta (2002), the wholesale prices reported by the retailer in the data are prices net of promotional payments such as off-invoice discounts and promotional moneys for feature advertisements and special displays, but not net of other sorts of promotional funds which the manufacturers could possibly make to the retailer. Therefore, if the manufacturers provide the retailer any of these "unobserved" promotional payments, marginal costs will be lower than the reported wholesale prices (or AAC's). Moreover, besides wholesales prices, the retailer may incur other expenses for specific operations such as shipping, storing, shelving, and pricing (Borin and Farris, 1990). If the retailer incorporates these expenses into marginal costs, MC's will be greater than wholesale prices (or AAC's).

Table 2.12 shows that almost all the ratios of MC's to AAC's of the NB's are less than 1, but those of the SB's are all greater than 1. A slow turnover of NB's inventories might cause their MC's to be less than the AAC's and the retailer's stockpiling of SB's during a trade promotion could cause their MC's to be greater than the AAC's. However, according to Chevalier et al (2000), DFF has good information on when a trade promotion is coming up and its optimal inventory management will incorporate information about such trade promotions relatively quickly into the AAC's. Moreover, according to Berges-Sennou et al (2003), due to the relative absence of barriers to entry in the food industry, private labels are often assumed to be produced by a competitive fringe composed of small firms. Therefore, in the literature, SB's are often assumed to be sold to retailers at marginal production costs, which implies that

SB manufacturers are unlikely to be able to afford trade promotions. Another explanation for NB's MC's being less than AAC's is that NB manufacturers may have made the retailer some sort of "unobserved" promotional payments which induces the retailer's extra support for their brands, but is not accounted for in the retailer's reported wholesale prices. On the other hand, since the retailer is the sole party who owns and controls SB's, all the expenses occurred to develop, maintain and promote SB's have to come out of the retailer's own pocket. This could be what causes SB's MC's to be greater than the AAC's.

2.5.5 Consumer surplus

Since the data consist of two sample periods: one when there are no SB's in the market and the other when SB's are available, using the demand-side estimates, I simulate consumer surplus before and after SB entry to examine whether the entry of SB's has a positive effect on consumer surplus. Note that only expected value of consumer surplus can be measured because the realization of the additive error term, ε_{ijt}, in the utility function is unobserved and thus needs to be integrated out. Moreover, since the data are market-level data, only the average expected consumer surplus can be measured by integrating over the expected consumer surplus of all consumers in a market. Equation (2.15) is the average expected consumer surplus per unit demand at market t. Equation (2.15) is approximated by Equation (2.16), i.e., the integral is approximated by taking 1000 random draws (ns=1000) from the distributions on demographics, $P^*(.)$. Equation (2.17) is the total expected consumer surplus at market t, derived by multiplying average expected consumer surplus by market size.

$$AECS_t(x, p_t, \delta_t; \theta_2, P^*) = \int_{D_i, \nu_i} ECS_{it} \, dP^*(D_i, \nu_i), \qquad \text{for } t = 1, \ldots, T \tag{2.15}$$

Where

$$ECS_{it}(x, p_t, \delta_t; \theta_2, D_i, \nu_i) = \frac{EU_{it}}{\alpha_i}$$

$$EU_{it} = \sum_{k=0}^{J_t} u_{ikt} \times Prob(\text{i chooses k}) = \ln[1 + \sum_{j=1}^{J_t} \exp(\tilde{u}_{ijt})]$$

$$u_{ijt} = \delta_{jt}(x_j, p_{jt}, \xi_j, \triangle\, \xi_{jt}; \theta_1) + \mu_{ijt}(x_j, p_{jt}, \nu_i, D_i; \theta_2) + \varepsilon_{ijt}$$

$$\tilde{u}_{ijt} = \delta_{jt}(x_j, p_{jt}, \xi_j, \triangle\, \xi_{jt}; \theta_1) + \mu_{ijt}(x_j, p_{jt}, \nu_i, D_i; \theta_2)$$

$$AECS^s_t = \frac{1}{ns} \sum_{i=1}^{ns} ECS_{it}, \qquad \text{for } t = 1, \ldots, T \tag{2.16}$$

$$TECS_t = AECS_t \times M_t, \qquad \text{for } t = 1, \ldots, T \tag{2.17}$$

Table 2.13 reports the means of average expected consumer surplus and total expected consumer surplus across all markets for the before and after SB entry periods. It is not too surprising that the average expected consumer surplus per unit demand is so small (only a little more than 1 cent) since the market share of the outside good is huge and the unit demand is defined as one serving per week. The monetary increase in average expected consumer surplus after SB entry seems trivial, but the percentage increase is pretty significant: consumer surplus is raised by over 10%. If we look at the total expected consumer surplus per market, after the entry of SB's, the monetary increase in consumer surplus is on average \$55.66 per store per week, still not very impressive. However, if this estimate can be applied to all of the stores owned by the supermarket chain (about 100 stores), the increase in consumer surplus is about \$5566 per week. Therefore, even though from an individual consumer's point of view, the entry of SB's does not seem to improve their welfare significantly, the aggregate benefit for the society as a whole can be quite substantial.

2.6 Conclusion

I estimate consumer demand for both national and store brands in a particular product category - hot-breakfast cereals - using scanner data from a multi-store supermarket chain in a major metropolitan area in generalized-method-of-moments estimation of a random-coefficients model. The empirical findings are: (1) After the entry of SB's, demand becomes more elastic for non-imitated NB's, and either more elastic or shows no change for imitated NB's; (2) in general, substitution patterns for NB's and SB's are asymmetric, i.e. when the prices of their favorite products increase, most NB buyers tend to substitute to other NB products, but SB buyers will substitute to the corresponding imitated NB's; (3) the increase in consumer surplus due to the entry of SB's is trivial for an individual consumer (\$0.001/week assuming that consumers have unit demand of one serving per week), but the aggregate benefit could be quite substantial (\$5566/week for the market under study).

In addition, I apply the demand-side estimates to the supply side, i.e., I back out the retailer's marginal costs assuming the retailer has constant and exogenous marginal costs and chooses prices to maximize category profits for each store. I suggested some explanations for the observed discrepancy between the estimated marginal costs and the observed average acquisition costs. The NB manufacturers might have provided the retailer with some promotional payments not revealed in the retailer's reported wholesale prices, and the retailer might have incurred expenses other than the wholesale costs in order to support and promote the SB's. Therefore, maintaining a successful SB program may require more resources from the retailers than simply paying the low wholesale prices.

The current model ignores possible strategic behaviors of the players, i.e., consumers, retailers, and manufacturers. For example, consumers may be forward-looking and they may wait for lower prices and stock up during a sale, which implies a dynamic demand model since consumers' purchase decisions will depend on current prices, their inventories, and their expectations of future prices. Moreover, there may exist strategic interactions between retailers' and manufacturers' pricing behaviors as well. These issues are certainly important and interesting though addressing them is beyond the scope of this study and is left for future research.

Chapter 3

A Reduced-Form Examination of Consumer Stockpiling Behavior in the Canned-Tuna Product Category

3.1 Introduction

For many consumer packaged goods, a temporary price reduction (a sale or price promotion) of a good usually causes a surge in demand for the good. In general, a sales bump can be decomposed into three components: brand switching, purchase incidence, and purchase quantity. If a good is storable, the effects of purchase incidence and purchase quantity can be further decomposed into two components: a consumption effect if consumers' consumption is price elastic and a stockpiling effect if consumers stock up for future usage. Composition of a sales bump may vary from good to good depending on factors such as storability, necessity, perceived differentiation, and purchase frequency. Hendel and Nevo (2002) find evidence of consumer inventory holding in the product categories of laundry detergents, soft drinks, and yogurt, and the relative strength of the consumption and stockpiling effects due to sales differs across the three product categories. Furthermore, using scanner data on laundry detergents, they estimate a structural dynamic demand model taking into account consumer stockpiling behavior and conclude that estimation derived from a standard static demand model is misleading if dynamics are present, but ignored (Hendel and Nevo, 2005).

I am interested in estimating consumer demand in the canned tuna product category. Since canned tuna is a storable good, when it is on sale, consumers can build up their inventories for future consumption. However, storability is only a necessary condition which enables stockpiling. When a good is non-staple, has low consumption frequency and is being promoted frequently, it is likely that sales may mostly prompt consumers to purchase for current consumption only. In other words, most consumers may decide not to stock up during sales to avoid storage cost. The question posed here then is whether consumer demand for canned tuna should be modeled by a dynamic model or a static one. On the one hand, estimating a dynamic model is computationally much more demanding compared to a static model. On the other hand, estimated price elasticities in a static model can be biased if dynamic motives are present, but ignored. I examine the data (including both household-level purchase data and store-level aggregate data) and run some reduced-form tests to investigate the significance of consumer inventory holding in the canned-tuna product category. The results suggest that only a limited amount of stockpiling exists in this product category. Therefore, the potential bias in the price elasticities estimated using a static demand model should not be a major concern.

3.2 Data

The data set I am using is known as "ERIM" provided by the James M. Kilts Center, GSB, University of Chicago. The ERIM data set was collected by the now-defunct ERIM division of A.C. Nielsen. It contains household-level purchase data on 9 packaged-good product categories in two mid-sized mid-western cities, Sioux Falls, SD and Springfield, MO. According to the data provider, these two cities are demographically typical of the U.S. population as a whole. Households who participated in the ERIM study were issued with magnetic ID cards. When they presented their ID cards at the checkout counters in the participating stores, their purchases of any UPC's in those 9 product categories were recorded. The total sales of the participating stores accounted for 80% of the market in terms of grocery and drug retail sales. Each record in the household-level data is a purchase occasion which records who made the purchase (household id), when (week) and where (store) the purchase occurred, the UPC and quantity being purchased, and the total payment made for the purchase. Moreover, when households joined the ERIM study, they were asked to fill out a questionnaire. Therefore, basic demographic information is available. In addition to household-level purchase data, ERIM also contains store-level aggregate data with information about total revenues, total quantities, feature advertisements, and in-store displays for each UPC in the 9 product categories on a weekly basis.

The product category studied here is canned tuna. I use both household- and store-level data in the city of Springfield, MO. The sample period is from January 1985 to May 1987 (123 weeks). 21 stores associated with 4 supermarket chains participated in the ERIM study in this market. The household-level data consist of 58,920 purchase occasions made by 5,255 households during the sample period. There are 5 package sizes in the canned-tuna product category: 3.25 oz, 6.5 oz, 9.25 oz, 12.5 oz, and 3-pack-3.25 oz. Table 3.1 presents the market share, purchase occasion share, and non-price promotion (i.e., feature advertisements and in-store displays) share of each package size. It shows that the 6.5 oz package size is dominant in terms of both store sales (89.6%) and household purchases (91.4%). Furthermore, Table 3.2 indicates that a handful of UPC's account for the majority of sales in this product category. They are 6.5-oz light tuna in water/oil of StarKist, Chicken of the Sea (CKN), Three Diamond, and private label (PL). The table shows the market, purchase-occasion, and non-price promotion shares and average shelf prices for these 4 brands. These UPC's have a collective market share of more than 79% and account for at least 81% of the households' purchase occasions and 95% of the participating stores' non-price promotional activities. Since consumers are likely to incur some storage cost for keeping inventories, it seems reasonable to assume that consumers do not stockpile when purchasing at regular (non-sale) prices. In other words, only when a good is on sale may the monetary saving give consumers incentives to stock up for future use. Therefore, frequent temporary price reductions are important in order to identify consumer stockpiling behavior. Note that except for the top sellers listed above, the prices of all other UPC's of canned tuna vary very little and their market shares are very small. Since the purpose is to investigate consumer stockpiling in the canned-tuna product category, I thus restrict the analyses to these top selling UPC's.

3.3 Descriptive Statistics

The store-level data provide information about each participating store's weekly total quantities, total revenues, and feature/dispaly activities for each UPC. I use the quantity and revenue variables to compute shelf prices for each brand-store-week combination. There are 4 brands, 21 stores, and 123 weeks. After excluding missing data, there are 9,407 brand-store-week observations. In general, retail prices, especially for StarKist and CKN, exhibit a certain pattern over time, i.e., the prices remain constant at a "regular" price for a relatively longer period of time and drop occasionally to a "sale" price for a shorter period. Figure 3.1 is an example of the retail price patterns of StarKist and CKN over the sample period in one participating store. Even though the kind of price patterns just described can be easily observed in the figure, it is harder to pinpoint the regular and sale prices. I follow one approach in Hendel and Nevo (2002), by which regular price is defined as the modal price over the sample period. I then define a brand to be on sale if its shelf price is at least 20% below its regular price. Table 3.3 presents the relation between aggregate demand and retail price discount. The table shows that boosts of sales due to price promotions are quite spectacular: about 78% of the quantity was sold during the sale weeks which accounted for about 25% of the time.

I also compare households' sale and non-sale purchase quantities. The sample consists of 1,781 households (who purchased only the 4 brands under study during the sample period) and 12,579 purchases. The lower panel of Table 3.4 shows that 79.7% of the purchase occasions occurred when the purchased brand was on sale. Therefore, it seems that price promotions are quite effective in inducing households to make a purchase. The upper panel of Table 3.4 reports the distributions of households' sale and non-sale purchase quantities. It seems that purchase quantity is more likely to be higher for a sale purchase. The percentage of non-sale purchases such that the quantity being purchased is no more than 4 cans is 99.1%, and 97.8% for sale purchases. Therefore, even though it is hard to say whether the increased quantities purchased during sales are used for current or future consumption, at least it shows that not many households stockpile a large number of cans during sales.

Table 3.5 presents the distribution of number of purchases for the 1,781 households and the distribution of purchase quantities for the 12,579 purchase occasions. The left panel of Table 3.4 shows that at least 50% of the households made no more than 5 purchases and less than 10% more than 15 during the sample period of 123 weeks. The right panel of Table 3.4 shows that less than 2% of the purchases involved a purchase quantity more than 4. Therefore, it seems that both purchase frequency and consumption of canned tuna are typically low.

Table 3.6 presents the averages of sale purchase ratio, purchase duration, and purchase quantity across households with the same purchase frequency. Column 4 ("Duration") shows that the lower the purchase frequency, the longer the average purchase duration, which is as expected if consumption of canned tuna is not serial correlated. Column 5 ("Quantity") shows that there is no clear correlation between purchase frequency and purchase quantity. Note that once opened, canned tuna becomes perishable and will be spoiled in a few days. Therefore, consumers are likely to consume canned tuna on a whole-can basis. In that sense, the small difference in the weekly purchase quantities among the households with different purchase frequencies may seem quite insignificant. Column 3 ("Ratio") shows that sales purchase ratios are generally very high for all the households, i.e., a household's purchases of canned tuna consist of mostly sales purchases regardless of his/her purchase frequency.

The descriptive statistics in this section suggest that most households are light and infrequent users of canned tuna and they made most of their purchases during sales. Therefore, it is possible that for most consumers price promotions of canned tuna only induce them to purchase for current consumption, but not much stock up for future use. In the next section, some reduced-form tests are provided to further investigate evidence of consumer stockpiling of canned tuna.

3.4 Reduced-Form Tests

The data directly observed by researchers only reveal consumers' purchase decisions, i.e., what and how many they buy. Since consumers' consumptions are not observable, so are their inventory levels. Therefore, even though researchers do observe that consumers purchase more during sales, it is not straightforward to separate a stockpiling effect from a consumption effect. Hendel and Nevo (2002) develop a dynamic model of consumer demand, and from that model they derive several predictions on the observed variables (regarding both household-level purchase patterns and store-level demand patterns) when consumer stockpiling behavior is present. Therefore, reduced-form tests based on those predictions may help to distinguish a dynamic model where both consumption and stockpiling effects are present during sales from a static one where only a consumption effect exists. The followings are their proposed testable predictions.

- *Prediction 1*: Duration to next purchase is longer for a sale purchase than a non-sale purchase.
- *Prediction 2*: Duration from previous purchase is shorter for a sale purchase than a non-sale purchase.
- *Prediction 3*: The probability that the previous purchase is non-sale is higher conditional on a non-sale purchase than a sale purchase.
- *Prediction 4*: Aggregate demand increases in duration since last sale. The duration effects are present during both sale and non-sale periods,

but stronger during the former.

In Hendel and Nevo's (2002) model of consumer inventory holding, a consumer's purchase decision follows a S-s type rule. s denotes the inventory level that triggers a purchase, which decreases in current price. Given that a purchase is made, S denotes the target inventory level, which also decreases in current price. According to the S-s decision rule, first, since a sale purchase results in a higher end of period inventory than a non-sale purchase, all else equal, the duration to the next purchase should be longer for a sale purchase; thus Prediction 1. Second, since a non-sale purchase is triggered by a lower beginning of period inventory than a sale purchase, on average, the duration from the previous purchase should be longer for a non-sale purchase; thus Prediction 2. Moreover, a lower inventory level triggers purchases at non-sale prices, which is more likely to occur if the previous purchase is non-sale; thus Prediction 3. Finally, since inventory decreases over time, demand should increase in duration from the previous sale. In addition, since the target inventory level is higher at sale prices than non-sale prices, the duration effect is stronger for the sale periods than the non sale periods.

Columns 1 and 2 in Table 3.7 present the average durations of sale and non-sale purchases to next purchase, respectively. Columns 4 and 5 present the average durations of sale and non-sale purchases from previous purchase, respectively. They are averages taken across purchases made by the households in the "household sample". The size of "household sample" is reduced gradually by excluding households with lower purchase frequencies. If consumers stockpile during sales, Column 1 should be greater than Column 2 and Column 4 should be smaller than Column 5. According to Column 3, the data are consistent with Prediction 1, i.e., duration to next purchase is longer for a sale purchase than a non-sale purchase. However, Column 6 shows that duration from previous purchase is longer for a sale purchase as well, which contradicts Prediction 2. The lower panel of Table 3.7 presents the within differences in duration to next and from previous purchases between sale and non-sale purchases, averaged out across the sample households. The results also show that durations to next purchase and from previous purchase are both longer for a sale purchase.

Columns 3 and 6 in Table 3.8 present the proportions of previous purchases being non-sale conditional on current purchases being sale and non-sale, respectively. Since Column 6 is greater than Column 3, the data are consistent with Prediction 3, i.e., previous purchases are more likely to be non-sale conditional on current purchases being non-sale. However, the lower panel of Table 3.8 shows that the within difference is positive, which contradicts Prediction 3.

Table 3.9 presents the regression results for duration effects on aggregate demand. This is a regression of log of quantity on log of price, duration since last sale, feature, display, Lent, store and brand dummies. The left panel of Table 3.9 shows that the coefficient on the duration variable is insignificant when sale and non-sale periods are pooled together in the regression. When I allow the duration effect during the sale periods to be different from that during the non-sale periods, the right panel of Table 3.9 shows that while the coefficient on duration since last sale is positive and significant during the sale periods, it is negative and insignificant during the non-sale periods. Therefore, the data are not entirely consistent with Prediction 4 which states that duration effects are present during both sale and non-sale periods.

Note that Hendel and Nevo's (2002) inventory model assumes that there is only one brand, i.e., brand choice is abstracted from the problem. However, in the data consumers do make brand choice since the product category consists of multiple brands. Suppose that there are two types of consumers in the sample: "brand-loyal" consumers and "switchers". The former only purchase a particular brand and the latter purchase whatever on sale. It is expected that "brand-loyal" consumers will stockpile their favorite brands during sales and thus display the kind of asymmetric duration patterns between sale and non-sale purchases as predicted by Hendel and Nevo's model. Judging from the descriptive statistics, it is likely that the frequency of sales (of any brand) is more often than that of consumption (of a typical consumer). In that case, there is no need for "switchers" to stockpile, and they only make sale purchases with symmetric purchase durations. If, say, "switchers" tend to have longer purchase durations than "brand-loyal" consumers, the data's being consistent with Prediction 1, but contrary to Prediction 2 may result from lumping purchases of both types of consumers together. One possible way to detect this problem is to run the reduced-form tests on a chain basis. Table 3.10 presents the frequencies of at least one brand on sale for each chain. Because there is variation in frequency of sales across chains, if consumer stockpiling behavior is present, the test results from the chains with lower frequencies of sales is more likely to be consistent with the predictions since the inventory effect should be stronger there. Tables 3.11 to 3.13 present the store-level statistics for the reduced-form tests.

Table 3.11 shows that for chains 1 to 3, both durations to the next purchase and from the previous purchase are longer for a sale purchase than a non-sale purchase, i.e., the chain-level test results are consistent with Prediction 1, but contradict Prediction 2, as before. As for chain 4, the results are consistent with both Predictions 1 and 2. However, it is not very convincing evidence of consumer stockpiling since chain 4 is actually the one chain with the most frequent sales of canned tuna. Table 3.12 shows that the chain-level tests for Prediction 3 are consistent with the prediction, as before. It should be expected since "switchers" make no non-sale purchases, i.e., the test statistics are based only on purchases by "brand-loyal" consumers. Finally, Table 3.13 shows that using a 5% significance level, (i) for chain 1, the duration effect is positive and significant during the sale periods, and positive and insignificant during the non-sale periods; (ii) for chain 2, the duration effect is positive and significant during both sale and non-sale periods;(iii) for chain 3, the duration effect is positive and significant during the sale periods, and negative and significant during the non-sale periods; and (iv) for chain 4, the duration effect is positive and insignificant during both sale periods and non-sale periods. Therefore, only the test result from chain 2 is entirely consistent with Prediction 4. However, since chains 1 and 3 have lower frequencies of sales than chain 2, we would expect the duration effects in chains 1 and 3 during the non-sale periods to be positive and significant as well.

According to the results of the above (non-chain-based and chain-based) reduced-form tests , there does not seem to be strong consumer stockpiling in the canned-tuna product category. Even though we cannot claim that no households stockpile canned tuna during sales, it is likely that the inventory effect is quite small and not very important in this product category (or "switchers" outnumber "brand-loyal" consumers).

3.5 Duration Model

In this section, I apply the data on household purchases to a duration model to study households' inter-purchase times of canned tuna. According to Jain and Vilcassim (1991), households' decisions about purchase timing may depend on marketing variables (such as price promotions, coupon offers, feature advertisements, and point of purchase displays), household characteristics, and the elapsed time since last purchase. To adjust for these factors' effects on the probability distribution of inter-purchase times, I use a proportional hazard function to analyze the duration data. The proportional hazard function has the form described by Equation (3.1).

$$h(t|X) = \lambda(t) \cdot \psi(X) \tag{3.1}$$

$\lambda(t)$ is the baseline hazard function which captures the effect of the elapsed time since last purchase. The parametric form of the baseline hazard function chosen is the expo-power (EP) hazard function which is very flexible in displaying a wide variety of shapes. The p.d.f. is

$$\lambda(t) = \gamma\alpha t^{\alpha-1}\exp(\theta t^{\alpha}),\ \gamma > 0,\ \alpha > 0 \tag{3.2}$$

γ, α, and θ are the parameters of the EP hazard function. Under alternative values of these parameters, the shape of the baseline hazard function can be flat, monotonically increasing, monotonically decreasing, U shaped or inverted U shaped.

$\psi(X)$ is a function of the covariates which shifts the hazard from its baseline. The parametric form of this function is

$$\psi(X) = \exp(X\beta) \tag{3.3}$$

The covariates, X, are a dummy variable indicating whether there is a sale when the purchase was made, a dummy variable indicating whether there is a feature advertisement when the purchase was made, a dummy variable indicating whether there is an in-store display when the purchase was made, a dummy variable indicating whether previous purchase is a sale purchase, household size, and log of household income. Note that consumer heterogeneity is accounted for through the observed household demographics only, i.e., I assume that unobserved household characteristics have no effects on the hazard rate. Moreover, I treat each spell of the inter-purchase times in the data as independent so that estimating this model is simple and fast.

Table 3.14 presents the estimation results of the proportional hazard model. Figure 3.2 displays the shape of the baseline hazard function. It is inverted U-shaped which implies that initially consumers are more likely to purchase canned tuna with the passage of time, but after a while without making any purchase, they eventually become less likely to do so. Regarding the estimated coefficients on the covariates, feature has a positive effect on the hazard rate, but is insignificant; the effects of display and household income on the hazard rate change from positive to negative and significant to insignificant as households with lower purchase frequencies are removed from the sample; and household size has a positive and significant effect on the hazard rate. Moreover, the covariate, previous purchase being a sale purchase, has a negative and significant effect on the hazard rate, which means that if a previous purchase is made during a sale period, then the inter-purchase time is expected to be longer. This seems to suggest consumer stockpiling behavior. However, the covariate, presence of a sale when the purchase is made, has a negative and significant effect on the hazard rate, which implies that price promotions decrease households' purchase probabilities. This does not make much sense and could be caused by the nature of the reduced-form specification. According to the descriptive statistics in Table 3.7, both durations from previous purchase and to next purchase are longer for a sale purchase than for a non-sale purchase. It is probably why when the duration data are used to estimate the hazard function of the duration model, the coefficients on the covariates of previous and current sale purchases are both negative, even though we would expect price promotions to accelerate households' purchase timings and thus shorten their inter-purchase times.

3.6 Conclusion

The results of the reduced-form tests for consumer inventory holding suggest that only a limited amount of stockpiling exists in the canned-tuna product category. Judging from the high ratio of sale purchases, the small amount of increased quantities of sale purchases, and the low purchase frequencies of the majority of households in the sample (as shown in Table 3.15), it is possible that most of the time, canned tuna are not on the households' shopping lists when they go grocery shopping. It is after they have arrived at a store and find canned tuna is on sale that they are prompted to purchase for current consumption. Since the inventory effect is small and not very important in this product category, it may be all right to estimate consumer demand for canned tuna using a static demand model under the assumption that consumers do not stockpile. The computational complexity can be greatly reduced and the potential bias in the price elasticities should not be serious.

Appendices

Appendix A

The MCMC Estimation Algotithm

Gibbs sampling as used in this paper consists of three layers, i.e. three conditional posterior distributions from which the model parameters (i.e., θ_i's, Π, and Σ_θ) are drawn. The first layer is for $\theta_i, i = 1, \ldots, N$; the conditional posterior of θ_i depends only on data for household i, rather than for the entire sample. The second and third layers are for Π and Σ_θ; their conditional posteriors do not depend on the data directly, but only on the draws of θ_i's, which themselves depend on the data. The details on how draws are taken are as follows.

1. Start with any initial values: θ_i^0's, Π^0, and Σ_θ^0.
2. For each $i, i = 1, \ldots, N$, draw θ_i^1 conditional on Π^0 and Σ_θ^0 using the Metropolis-Hastings algorithm which operates as follows:
   (a) Draw a $k \times 1$ vector υ_i^0 from a standard normal density.
   (b) Create a trial value $\tilde{\theta}_i^1$: $\tilde{\theta}_i^1 = \theta_i^0 + \pi L^0 \upsilon_i^0$. Here L^0 is the Choleski factor of Σ_θ^0, i.e. $L^0 L^{0'} = (\Sigma_\theta^0)^{-1}$and π is a scalar which determines the size of each jump. Usually, smaller (larger) jumps result in more (fewer) accepts, and thus the MH algorithm takes more (less) iterations to converge and the series of draws are more (less) correlated after convergence. The value of π is programmed to vary at each iteration to achieve an acceptance rate of about 0.3.

(c) Draw a variable u_i^1 from a standard uniform density and calculate the ratio F_i^1:

$$F_i^1 = \frac{L(y_i|\tilde{\theta}_i^1)\phi(\tilde{\theta}_i^1|\Pi^0, \Sigma_\theta^0)}{L(y_i|\theta_i^0)\phi(\theta_i^0|\Pi^0, \Sigma_\theta^0)}$$

If $u_i^1 \leqslant F_i^1$, accept $\tilde{\theta}_i^1$ and let $\theta_i^1 = \tilde{\theta}_i^1$; if $u_i^1 > F_i^1$, reject $\tilde{\theta}_i^1$ and let $\theta_i^1 = \theta_i^0$.

3. Draw Π^1 conditional on θ_i^1's and Σ_θ^0 from its normal conditional posterior which has the following form:

$$vec(\Pi^1) \sim N(vec(\Phi^1), \Sigma_\theta^0 \otimes (D'D + I)^{-1})$$

where

$$\Phi^1 = (D'D + I)^{-1}(D'\Theta^1), \underset{(N\times k)}{\Theta^1} = \begin{pmatrix} \theta_i^{1'} \\ \vdots \\ \theta_N^{1'} \end{pmatrix}, \underset{(N\times k)}{D} = \begin{pmatrix} D_1' \\ \vdots \\ D_N' \end{pmatrix}$$

Draws from the above multivariate normal distribution are obtained as follows:

(a) Draw a $kd \times 1$ vector v^0 from the standard normal density.

(b) Calculate the new draw Π^1: $vec(\Pi^1) = vec(\Phi^1) + L_\Pi^0 v^0$. Here L_Π^0 is the Choleski factor of $(\Sigma_\theta^0 \otimes (D'D + I)^{-1})$.

4. Draw Σ_θ^1 conditional on θ_i^1's and Π^1 from its inverted Wishart conditional posterior which has the following form:

$$\Sigma_\theta^1 \sim IW(v_1, \frac{v_0 I + N s^1}{v_0 + N})$$

$$v_1 = v_0 + N, s^1 = diag\left(\frac{1}{N}\sum_{i=1}^{N}(\theta_i^1 - \Pi^1 D_i)^2\right)$$

Draws from the above inverted Wishart are obtained as follows:

(a) Draw v_1 $k \times 1$ vectors from the standard normal density. Label these draws by $\tau_m^0, m = 1, \dots, v_1$.

(b) Calculate the new draw Σ_θ^1: $\Sigma_\theta^1 = (R^1)^{-1}$ where

$$R^1 = \frac{1}{v_1}\sum_{m=1}^{v_1}(L_\Sigma^0 \tau_m^0)(L_\Sigma^0 \tau_m^0)', L_\Sigma^0 L_\Sigma^{0\prime} = (\frac{v_0 I + N s^1}{v_0 + N})^{-1}$$

Repeat the last three steps many times. Iterating through numerous cycles of draws from the conditional posteriors eventually provides draws from the joint posterior of Π, Σ_θ, and θ_i's. The mean and standard deviation of the draws retained after convergence has been achieved can then be calculated to obtain the estimates and standard errors of the parameters

When the variance σ^2 of a normal distribution is unknown, it is often assumed that the prior of σ^2 is inverted gamma distribution. If X is inverted gamma distributed with degrees of freedom v and scale s, its probability density function has the following form:

$$f(x) = \frac{\exp(-vs/2x)}{cx^{(v+1)/2}}$$

where c is a normalizing constant. The reasons for choosing the inverted gamma are (i) its density is zero for any negative value of σ^2, reflecting the fact a variance must be positive and (ii) under the inverted gamma prior, the posterior of σ^2 is also inverted gamma. An inverted Wishart distribution is the multivariate generalization of an inverted gamma distribution.

Appendix B

Tables

Table 1.1: Shares of different package sizes

package size	market share	purchase occasion share	feature/display share
3.25 oz	0.22%	0.18%	0.13%
6.5 oz	89.58%	91.40%	99.48%
9.25 oz	4.45%	3.82%	0.39%
12.5 oz	4.20%	2.97%	0.00%
3-pack-3.25 oz	1.55%	1.63%	0.00%
total	100%	100%	100%

Note: 1. "market share" is the quantity (measured by oz) sold of a particular package size over the total quantity sold of the whole product category. 2. "purchase occasion share" is the number of purchase occasions of a particular package size over the total number of purchase occasions. 3. "feature/display share" is the number of weeks of a particular package size on feature/display over the sum of the number of weeks of each package size on feature/display.

Table 1.2: Shares and prices of the dominant brands

brand	market share	purchase occasion share	feature/display share	avg shelf price
1: StarKist	42.17%	45.28%	56.94%	0.72
2: CKN	26.42%	26.19%	26.66%	0.78
3: 3 Diamond	2.18%	2.75%	6.39%	0.75
4: PL	8.37%	7.37%	5.96%	0.67
total	79.15%	81.59%	95.95%	

Note: "avg shelf price" is the average retail price of a particular brand across stores and sample weeks and is measured by dollar.

Table 1.3: Frequencies of sales of a brand

On Sale	Chain	StarKist	CKN	3 Diamond	PL
disc >= 20%	Chain 1	24.39%	8.94%	0.00%	15.09%
	Chain 2	37.36%	31.32%	0.00%	0.89%
	Chain 3	35.77%	18.21%	8.13%	0.81%
	Chain 4	70.73%	71.34%	21.65%	10.06%
disc >= 30%	Chain 1	7.32%	8.67%	0.00%	14.54%
	Chain 2	13.42%	8.50%	0.00%	0.00%
	Chain 3	29.59%	16.26%	4.07%	0.00%
	Chain 4	44.51%	51.83%	0.00%	0.00%
disc >= 40%	Chain 1	3.79%	5.15%	0.00%	2.35%
	Chain 2	7.16%	2.46%	0.00%	0.00%
	Chain 3	29.43%	12.68%	4.07%	0.00%
	Chain 4	14.63%	24.70%	0.00%	0.00%

Note: A brand is defined to be on sale if $disc = \dfrac{shelf - regular}{regular} \geq d\%$

Table 1.4: Frequencies of the number of brands on sale

Sales	Chain	no brand	1 brand	2 brands	3 brands	4 brands
disc>=20%	Chain 1	57.72%	36.31%	5.78%	0.18%	0.00%
	Chain 2	45.41%	39.60%	14.99%	0.00%	0.00%
	Chain 3	52.20%	33.50%	13.50%	0.81%	0.00%
	Chain 4	1.83%	34.76%	51.83%	10.98%	0.61%
disc>=30%	Chain 1	74.71%	20.23%	4.88%	0.18%	0.00%
	Chain 2	78.75%	20.58%	0.67%	0.00%	0.00%
	Chain 3	62.60%	25.69%	10.89%	0.81%	0.00%
	Chain 4	28.05%	47.56%	24.39%	0.00%	0.00%
disc>=40%	Chain 1	89.07%	10.66%	0.18%	0.09%	0.00%
	Chain 2	91.05%	8.28%	0.67%	0.00%	0.00%
	Chain 3	66.34%	21.95%	10.89%	0.81%	0.00%
	Chain 4	65.24%	30.18%	4.57%	0.00%	0.00%

Table 1.5: The MCMC estimates of the population-level parameters

	Π			Σ_q
	constant	**income**	**size**	
StarKist	3.70710**	0.55210	0.22897	0.12875**
	(1.27206)	(0.61372)	(0.17600)	(0.02705)
CKN	3.34931**	0.59425	0.24154	0.12078**
	(1.27926)	(0.61521)	(0.17530)	(0.02900)
3 Diamond	2.90502**	0.39606	0.18709	0.53709**
	(1.30152)	(0.60893)	(0.18160)	(0.16570)
Private Label	2.68617**	0.42876	0.20891	0.50156**
	(1.29626)	(0.61165)	(0.17601)	(0.10804)
No Purchase	5.19037**	0.75619	0.10827	0.20870**
	(1.26780)	(0.60572)	(0.17990)	(0.04209)
Price	-2.64889**	0.27691**	-0.05549	0.18070**
	(0.21832)	(0.07134)	(0.03765)	(0.03033)
Advertising	1.10652**	-0.10217**	-0.04004	0.16352**
	(0.17815)	(0.05835)	(0.03249)	(0.03157)
Lapse	0.06511**	-0.01249**	0.00274	0.00023
	(0.01771)	(0.00669)	(0.00320)	(0.00017)
Lapse^2	-0.00116**	0.00019	-0.00003	0.00000
	(0.00038)	(0.00014)	(0.00006)	(0.00000)
Alpha*	-1.93368**	-0.12656	-0.01063	0.50265**
	(0.40801)	(0.15018)	(0.06747)	(0.10601)
Rho*	-0.01109	-0.21276**	0.07076	0.50957**
	(0.39637)	(0.12617)	(0.07234)	(0.21466)

Note: 1. A double asterisk indicates that the estimates are statistically significant at a 10% level.

2. Standard error is in the parenthesis.

Table 1.6: Proportions of consumer heterogeneity explained by the demographic variables

Individual-level parameter	l^2
g_1 **(StarKist)**	66.85%
g_2 **(CKN)**	70.32%
g_3 **(3 Diamond)**	34.45%
g_4 **(Private Label)**	36.54%
g_0 **(No Purchase)**	57.39%
g_P **(Price)**	16.63%
g_A **(Advertising)**	6.32%
b_1 **(Lapse)**	36.25%
b_2 **(Lapse^2)**	30.00%
a	3.66%
r	6.07%

Table 1.7: The own- and cross-price elasticities of each brand and the elasticities of no-purchase probability

	StarKist	CKN	3 Diamond	PL
StarKist	-3.3895	0.8294	0.0913	0.1570
	(0.9809)	(0.4976)	(0.0848)	(0.1219)
CKN	1.5337	-4.2362	0.0922	0.1552
	(0.4951)	(0.7745)	(0.0785)	(0.1164)
3 Diamond	1.2666	0.6892	-4.4174	0.1738
	(0.4351)	(0.4173)	(0.5103)	(0.1405)
Private Label	1.4925	0.7940	0.1107	-4.4674
	(0.5418)	(0.5077)	(0.1111)	(0.5344)
No Purchase	0.1089	0.0506	0.0061	0.0093
	(0.0619)	(0.0411)	(0.0066)	(0.0073)

Note: 1. The entry in cell (i, j) gives the percentage change in brand *i*'s quantity (in the outside good's probability) with respect to the percentage change in brand *j*'s price. 2. Standard error is in the parenthesis.

Table 1.8: Sales bump decomposition and percentage changes in demand

10% price discount	StarKist	CKN	3 Diamond	PL	average
Sales bump decomposition					
brand switching	24.46%	48.66%	51.16%	50.29%	43.64%
purchase incidence	58.11%	41.02%	37.96%	34.63%	42.93%
purchase quantity	17.43%	10.32%	10.88%	15.08%	13.43%
Percentage demand change					
promoted brand	37.35%	57.03%	60.99%	61.53%	54.23%
product category	18.63%	7.57%	0.91%	1.55%	7.16%
20% price discount	**StarKist**	**CKN**	**3 Diamond**	**PL**	**average**
Sales bump decomposition					
brand switching	19.75%	44.31%	49.72%	48.04%	40.46%
purchase incidence	62.61%	43.77%	37.55%	35.43%	44.84%
purchase quantity	17.63%	11.92%	12.73%	16.53%	14.70%
Percentage demand change					
promoted brand	86.97%	150.62%	175.49%	169.72%	145.70%
product category	46.08%	21.69%	2.69%	4.47%	18.73%
30% price discount	**StarKist**	**CKN**	**3 Diamond**	**PL**	**average**
Sales bump decomposition					
brand switching	15.68%	38.09%	47.75%	44.92%	36.61%
purchase incidence	67.43%	48.34%	36.95%	36.96%	47.42%
purchase quantity	16.89%	13.56%	15.30%	18.13%	15.97%
Percentage demand change					
promoted brand	152.85%	283.80%	380.84%	354.21%	292.93%
product category	85.10%	45.43%	6.07%	9.88%	36.62%
40% price discount	**StarKist**	**CKN**	**3 Diamond**	**PL**	**average**
Sales bump decomposition					
brand switching	12.24%	31.26%	44.80%	40.77%	32.27%
purchase incidence	71.97%	54.48%	37.64%	39.28%	50.84%
purchase quantity	15.80%	14.26%	17.56%	19.95%	16.89%
Percentage demand change					
promoted brand	240.73%	467.01%	766.71%	640.52%	528.74%
product category	139.50%	83.01%	12.90%	19.22%	63.66%

50% price discount	StarKist	CKN	3 Diamond	PL	average
Sales bump decomposition					
brand switching	9.44%	24.54%	40.31%	35.94%	27.56%
purchase incidence	76.50%	61.31%	39.95%	43.13%	55.22%
purchase quantity	14.06%	14.15%	19.74%	20.94%	17.22%
Percentage demand change					
promoted brand	362.71%	724.58%	1406.39%	1085.70%	894.84%
product category	216.88%	141.38%	25.59%	35.23%	104.77%

Table 1.9: Average estimated retailer marginal costs

Retailer MC ($)	Chain 1	Chain 2	Chain 3	Chain 4
StarKist	0.46	0.41	0.27	0.44
	(0.100)	(0.146)	(0.218)	(0.107)
CKN	0.47	0.44	0.41	0.42
	(0.122)	(0.121)	(0.164)	(0.130)
3 Diamond	0.44	0.44	0.35	0.41
	(0.049)	(0.038)	(0.094)	(0.062)
PL	0.39	0.33	0.33	0.33
	(0.092)	(0.029)	(0.053)	(0.039)

Note: Standard error is in the parenthesis.

Table 1.10: Retailer pass-through for StarKist and CKN

~ 20% retail price di scount		price promotion		
	none	both	StarKist	CKN
Retail Price ($)				
StarKist	0.89	0.69	0.69	0.89
CKN	0.89	0.69	0.89	0.69
Wholesale Price ($)				
StarKist	0.54	0.38	0.39	0.54
CKN	0.54	0.38	0.53	0.39
Retailer Pass-through				
StarKist		124.6%	127.2%	
CKN		125.2%		130.4%
~ 30% retail price discount		**price promotion**		
	none	**both**	**StarKist**	**CKN**
Retail Price ($)				
StarKist	0.89	0.59	0.59	0.89
CKN	0.89	0.59	0.89	0.59
Wholesale Price ($)				
StarKist	0.54	0.30	0.31	0.53
CKN	0.54	0.30	0.53	0.31
Retailer Pass-through				
StarKist		123.1%	125.5%	
CKN		123.8%		128.6%
~ 40% retail price discount		**price promotion**		
	none	**both**	**StarKist**	**CKN**
Retail Price ($)				
StarKist	0.89	0.49	0.49	0.89
CKN	0.89	0.49	0.89	0.49
Wholesale Price ($)				
StarKist	0.54	0.21	0.22	0.52
CKN	0.54	0.21	0.51	0.23
Retailer Pass-through				
StarKist		121.1%	123.4%	
CKN		121.8%		126.4%

Table 1.11: Estimated losses of category profits from retailers' sales

WP	Optimal-StarKist	Optimal-CKN	Demand-StarKist	Demand-CKN	Profit - Scenario 1	Profit - Scenario 2	Ratio
0.49	0.82	0.82	31.40	18.55	30.03	33.18	90.51%
0.50	0.83	0.84	30.23	17.87	28.79	32.20	89.40%
0.51	0.85	0.85	29.11	17.22	27.54	31.26	88.12%
0.52	0.86	0.86	28.04	16.60	26.30	30.35	86.66%
0.53	0.87	0.88	27.01	16.00	25.05	29.47	85.02%
0.54	**0.88**	**0.89**	**26.03**	**15.43**	**23.81**	**28.63**	**83.18%**
0.55	0.90	0.90	25.09	14.88	22.57	27.81	81.14%
0.56	0.91	0.91	24.19	14.36	21.32	27.03	78.89%
0.57	0.92	0.93	23.33	13.86	20.08	26.27	76.43%
0.58	0.93	0.94	22.51	13.38	18.83	25.54	73.74%
0.59	0.95	0.95	21.72	12.92	17.59	24.83	70.82%
0.60	0.96	0.96	20.96	12.48	16.34	24.15	67.67%
0.61	0.97	0.98	20.23	12.05	15.10	23.50	64.26%
0.62	0.98	0.99	19.54	11.65	13.85	22.86	60.61%
0.63	1.00	1.00	18.87	11.26	12.61	22.25	56.68%
0.64	1.01	1.01	18.23	10.89	11.37	21.66	52.49%

	no sale	sale-StarKist	sale-CKN
Retail Price-StarKist	0.89	0.69	0.89
Retail Price- CKN	0.89	0.89	0.69
Quantity-StarKist	25.42	57.10	16.68
Quantity-CKN	15.47	9.04	41.62

Note: "Ratio" is calculated as "Profit– Scenario 1" over "Profit– Scenario 2".

Table 1.12: Estimated losses from retailers' cyclical pricing

city	Store #	zip	inc(mean)	inc(std)	ratio
OAK LAWN	8	60453	10.8050	0.7483	68.26%
RIVER GROVE	18	60171	10.6410	0.7296	72.62%
PARK RIDGE	32	60068	11.2120	0.7305	77.38%
NORTHBROOK	52	60062	11.3950	0.8604	73.43%
CHICAGO	53	60662	10.6420	0.8629	80.02%
CRYSTAL LAKE	59	60014	11.1620	0.6237	56.11%
JOLIET	70	60435	10.7490	0.7235	61.83%
NORTH RIVERSIDE	71	60546	10.9250	0.7529	75.91%
CHICAGO	73	60629	10.6310	0.8173	70.37%
DOWNERS GROVE	78	60516	11.2280	0.6241	60.11%
ARLINGTON HEIGHTS	80	60005	10.9840	0.7762	62.16%
LANSING	83	60438	10.8030	0.6783	62.79%
EVANSTON	93	60202	10.9620	0.7564	79.47%
CHICAGO	98	60638	10.7420	0.8073	72.81%
CHICAGO	100	60608	10.2020	0.9323	81.10%
MERRIONETTE PARK	102	60655	11.0390	0.6096	68.14%

Note: 1. "inc(mean)" and "inc(std)" are the two estimated parameters of the lognormal distribution which household income is assumed to follow; household income is measured by dollars. 2. "ratio" is calculated as the 'total data profit' over the 'total optimal profit'; see text in Section 1.8.

Table 2.1: List of UPCs in the demand system before aggregation

	UPC	description	product	weight (oz)	serving
1	1313000612	Nabisco Cream of Wheat 2.5 MIN	1	28	24
2	1313000622	Nabisco Cream of Wheat 1 MIN	1	28	24
3	1313006025	Nabisco Inst Wheat-Regular	2	12	12
4	1313006057	Nabisco Inst Wheat-Variety	3	12.5	10
5	3000001020	Quaker Old Fashioned Oats	4	18	13
6	3000001040	Quaker Old Fashioned Oats	5	42	30
7	3000001180	Quaker Quick Oats	4	18	13
8	3000001200	Quaker Quick Oats	5	42	30
9	3000001210	Quaker Inst Oats-Regular	6	12	12
10	3000001190	Quaker Inst Oats-MP	7	15	10
11	3000001240	Quaker Inst Oats-RS	7	13	10
12	3000001340	Quaker Inst Oats-CINNA	7	12.5	10
13	3000001380	Quaker Inst Oats-APPLE	7	12.5	10
14	3000001460	Quaker Inst Oats-Variety	7	13.3	10
15	3000001800	Quaker Inst Oats-PEACH	7	12.5	10
16	3000001820	Quaker Inst Oats-STRAW	7	12.5	10
17	3000001880	Quaker Inst Oats-FRUIT	7	12.5	10
18	3828125073	Dominick Quick Oats	8	18	13
19	3828125077	Dominick Quick Oats	9	42	30
20	3828125081	Dominick Inst Oats-Regular	10	12	12
21	3828125085	Dominick Inst Oats-MP	11	15	10
22	3828125098	Dominick Inst Oats-Variety	11	13	10
23	3828125105	Dominick Inst Oats-APPLE	11	12.3	10

Table 2.2: List of products in the demand system and their observed product characteristics

product	description	const	wheat/Nab	inst/SP	big	flavor	SB
1	Nabisco Non-Inst Wheat	1	1	0	0	0	0
2	Nabisco Inst Wheat-Regular	1	1	1	0	0	0
3	Nabisco Inst Wheat-Flavored	1	1	1	0	1	0
4	Quaker Non-Inst Oats	1	0	0	0	0	0
5	Quaker Non-Inst Oats- Big Tube	1	0	0	1	0	0
6	Quaker Inst Oats-Regular	1	0	1	0	0	0
7	Quaker Inst Oats-Flavored	1	0	1	0	1	0
8	Dominick Non-Inst Oats	1	0	0	0	0	1
9	Dominick Non-Inst Oats-Big Tube	1	0	0	1	0	1
10	Dominick Inst Oats-Regular	1	0	1	0	0	1
11	Dominick Inst Oats-Flavored	1	0	1	0	1	1

Table 2.3: List of stores and store-level demographics

	store	city	zip code	mean (loginc)	sigma (loginc)	hhund18	hhsize	avg traffic
1	DOMINICKS 8	OAK LAWN	60453	10.81	0.75	27.9%	2.46	23208
2	DOMINICKS 18	RIVER GROVE	60171	10.64	0.73	30.9%	2.44	21719
3	DOMINICKS 32	PARK RIDGE	60068	11.21	0.73	33.6%	2.61	26841
4	DOMINICKS 52	NORTHBROOK	60062	11.40	0.86	35.1%	2.63	18137
5	DOMINICKS 53	CHICAGO	60662	10.64	0.86	34.6%	2.69	13734
6	DOMINICKS 59	CRYSTAL LAKE	60014	11.16	0.62	46.6%	2.93	12584
7	DOMINICKS 70	JOLIET	60435	10.75	0.72	32.6%	2.43	18562
8	DOMINICKS 71	NORTH RIVERSIDE	60546	10.93	0.75	27.1%	2.34	20103
9	DOMINICKS 73	CHICAGO	60629	10.63	0.82	52.5%	3.51	22452
10	DOMINICKS 78	DOWNERS GROVE	60516	11.23	0.62	37.8%	2.73	16631
11	DOMINICKS 80	ARLINGTON HEIGHTS	60005	10.98	0.78	26.3%	2.3	20085
12	DOMINICKS 83	LANSING	60438	10.80	0.68	32.8%	2.49	20158
13	DOMINICKS 93	EVANSTON	60202	10.96	0.76	30.2%	2.3	18634
14	DOMINICKS 98	CHICAGO	60638	10.74	0.81	32.4%	2.69	24541
15	DOMINICKS 100	CHICAGO	60608	10.20	0.93	49.1%	3.31	24074
16	DOMINICKS 102	MERRIONETTE PARK	60655	11.04	0.61	35.5%	2.67	24463

Note: 1. “mean” and “sigma” are the two estimated parameters of the lognormal distribution which household income is assumed to follow. 2. Household income is measured by dollars. 3. “avg traffic” is the weekly number of households visiting the store averaged over the sample period.

Table 2.4: Descriptive statistics of the data

	quantity				retail profit			
	Before	After	change	%C	Before	After	change	%C
Product 1	466	336	-130	-27.89	20.42	23.48	3.06	14.98
Product 2	134	145	11	8.53	20.31	22.45	2.13	10.51
Product 3	76	50	-26	-34.09	20.64	22.88	2.24	10.87
Product 4	1269	1057	-212	-16.7	18.86	20.5	1.64	8.68
Product 5	1226	1046	-180	-14.66	18.64	21.66	3.01	16.17
Product 6	315	223	-91	-29.01	19.95	24.76	4.81	24.09
Product 7	1235	935	-301	-24.34	20.3	23.47	3.17	15.62
Product 8	—	184	—	—	—	55.84	—	—
Product 9	—	295	—	—	—	53.67	—	—
Product 10	—	91	—	—	—	52.21	—	—
Product 11	—	277	—	—	—	52.16	—	—

	price				aac			
	Before	After	change	%C	Before	After	change	%C
Product 1	11.74	11.97	0.24	2.02	9.29	9.15	-0.14	-1.47
Product 2	22.11	23.78	1.67	7.57	17.52	18.41	0.9	5.11
Product 3	26.62	28.52	1.91	7.16	21.01	21.95	0.94	4.48
Product 4	13.67	12.5	-1.18	-8.6	10.98	9.88	-1.1	-10.06
Product 5	10.38	9.51	-0.87	-8.39	8.39	7.43	-0.96	-11.39
Product 6	22.29	22.22	-0.07	-0.33	17.62	16.66	-0.96	-5.45
Product 7	26.73	26.61	-0.13	-0.48	21.03	20.3	-0.73	-3.47
Product 8	—	9.62	—	—	—	4.12	—	—
Product 9	—	7.74	—	—	—	3.54	—	—
Product 10	—	16.88	—	—	—	7.82	—	—
Product 11	—	20.24	—	—	—	9.38	—	—

product	price			aac			profit			quantity		
	NB	SB	ratio	NB	SB	ratio	NB	SB	ratio	NB	SB	ratio
4 vs 8	12.5	9.62	0.77	9.88	4.12	0.42	20.5	55.84	2.72	1046	184	0.18
5 vs 9	9.51	7.74	0.81	7.43	3.54	0.48	21.66	53.67	2.48	1464	295	0.2
6 vs 10	22.22	16.88	0.76	16.66	7.82	0.47	24.76	52.21	2.11	223	91	0.41
7 vs 11	26.61	20.24	0.76	20.3	9.38	0.46	23.47	52.16	2.22	935	277	0.3

Note: 1. "quantity" is the average weekly number of servings sold. 2. "price" and "aac" are the average retail price and average aac per serving measured by cents. 3. "profit" is calculated by (price-acc)/price*100. 4. "change" is calculated by (After-Before) and "%C" is calculated by (change/Before*100). 5. "ratio" is calculated by (SB / NB).

Table 2.5: Estimation results for the logit model

Set A	(i)	(ii)	Set B	(iii)	(iv)
Covariate	**OLS**	**2SLS**	**Covariate**	**OLS**	**2SLS**
Price	-11.636	-18.919	**Price**	-8.882	-6.367
	(0.2106)	(0.5566)		(0.1844)	(1.1189)
Summer	-0.452	-0.447	**Summer**	-0.463	-0.465
	(0.0092)	(0.0094)		(0.0074)	(0.0075)
Promotion	0.418	0.239	**Promotion**	0.417	0.477
	(0.0133)	(0.0186)		(0.0108)	(0.0284)
Constant	-2.245	-1.294	**Product 1**	-3.748	-4.049
	(0.0292)	(0.0734)		(0.0242)	(0.1341)
Wheat	-1.441	-1.446	**Product 2**	-3.930	-4.509
	(0.0098)	(0.0100)		(0.0437)	(0.2580)
Inst/SP	-0.084	0.619	**Product 3**	-4.343	-5.037
	(0.0231)	(0.0549)		(0.0519)	(0.3090)
Big	-0.231	-0.405	**Product 4**	-2.696	-3.036
	(0.0144)	(0.0191)		(0.0269)	(0.1517)
Flavor	0.903	1.223	**Product 5**	-2.816	-3.069
	(0.0144)	(0.0269)		(0.0210)	(0.1130)
SB	-2.070	-2.376	**Product 6**	-3.324	-3.892
	(0.0143)	(0.0261)		(0.0429)	(0.2533)
			Product 7	-1.613	-2.299
				(0.0513)	(0.3051)
			Product 8	-4.986	-5.240
				(0.0242)	(0.1140)
			Product 9	-4.592	-4.792
				(0.0211)	(0.0899)
			Product 10	-5.052	-5.495
				(0.0360)	(0.1977)
			Product 11	-3.614	-4.142
				(0.0418)	(0.2356)
Adj R^2	0.692	na	**Adj R^2**	0.803	na
J	na	819	**J**	na	1451

Note: The number in the parenthesis is the standard error of the estimate.

Table 2.6: Median own-price elasticities in the logit model

	Before	After	Change	%Change
Product 1	0.675	0.7573	0.0823	12.1919
Product 2	1.2897	1.5175	0.2278	17.661
Product 3	1.5484	1.8228	0.2744	17.7194
Product 4	0.772	0.797	0.0251	3.2467
Product 5	0.5831	0.6013	0.0181	3.1123
Product 6	1.3285	1.4505	0.122	9.1846
Product 7	1.5617	1.7207	0.159	10.1813
Product 8	—	0.651	—	—
Product 9	—	0.5001	—	—
Product 10	—	1.1488	—	—
Product 11	—	1.376	—	—

Note: "%Change" is calculated by (After-Before)/Before*100.

Table 2.7: Estimation results for the full model

covariate	mean	std. dev.	hhund18	logincome
price	-113.717*	2.174*	1.579	4.908*
	(2.1250)	(0.2925)	(8.2903)	(0.3834)
constant	-102.680*	2.437*	1.164	9.066*
	(0.1647)	(0.0696)	(1.0608)	(0.0472)
wheat/Nab	30.375*	1.268*	-2.185*	-2.693*
	(0.0278)	(0.1445)	(0.1557)	(0.0116)
inst/SP	-28.616*	3.398*	0.818	2.442*
	(0.1187)	(0.0760)	(0.8558)	(0.0421)
big	13.451*	1.792*	-5.212*	-1.180*
	(0.0461)	(0.2330)	(1.8633)	(0.0156)
flavor	-83.645*	2.825*	13.295*	6.433*
	(0.0694)	(0.1248)	(2.2613)	(0.1925)
SB	34.804*	2.530*	-1.861*	-3.364*
	(0.0584)	(0.2669)	(0.4124)	(0.0314)
summer	-1.777*	—	—	—
	(0.0135)	—	—	—
promotion	0.167*	—	—	—
	(0.0267)	—	—	—

Note: 1. Product-, store-, and year- specific dummy variables are included as covariates in the regressions. 2. A single asterisk denotes that the estimates are significant at a 5% significance level. 3. Standard errors of the estimates are reported in the parentheses.

Table 2.8: Median own-price elasticities and 95% confidence intervals in the full model

	Before			After			% Change		
	Point Est.	[95% confidence interval]		Point Est.	[95% confidence interval]		Point Est.	[95% confidence interval]	
Product 1	5.850	5.318	6.468	6.238	5.589	6.905	6.641	4.396	7.037
Product 2	11.510	10.203	12.768	12.656	10.969	13.758	9.957	6.325	9.420
Product 3	13.760	11.858	15.418	15.265	13.187	17.055	10.941	9.147	11.075
Product 4	5.320	4.948	5.662	5.378	4.957	5.641	1.092	-1.997	1.908
Product 5	3.750	3.310	4.463	3.792	3.334	4.446	1.123	-3.507	3.380
Product 6	10.930	10.363	11.327	11.491	10.810	11.914	5.139	2.535	6.047
Product 7	4.873	4.214	5.659	7.204	6.193	8.264	47.831	28.929	32.959
Product 8	—	—	—	5.428	4.999	5.754	—	—	—
Product 9	—	—	—	4.164	3.566	4.814	—	—	—
Product 10	—	—	—	9.532	8.453	10.053	—	—	—
Product 11	—	—	—	10.082	8.691	11.273	—	—	—

Note: 1. Own-price elasticities are reported in absolute value. 2. “% Change” is calculated by (after-before)/before*100.

Table 2.9: Median own- and cross-price elasticities in the full model

Before	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	Product 7
Product 1	-5.8499	0.1333	0.0146	0.9607	0.6364	0.2024	0.1213
Product 2	0.2594	-11.5098	0.0392	0.3875	0.4305	0.9198	0.3841
Product 3	0.0392	0.0618	-13.7597	0.2726	0.0052	0.4163	8.0851
Product 4	0.2978	0.0693	0.0331	-5.3201	0.4541	0.3448	0.9567
Product 5	0.2525	0.0833	0.0008	0.5454	-3.7497	0.1760	0.0240
Product 6	0.1605	0.3884	0.1059	0.8408	0.3508	-10.9297	1.4976
Product 7	0.0208	0.0381	0.5212	0.5032	0.0100	0.2960	-4.8731
outside	0.0306	0.0145	0.0032	0.0651	0.0657	0.0226	0.0456
variation	14	27	635	15	122	41	337

After	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	Product 7	Product 8	Product 9	Product 10	Product 11
Product 1	-6.2384	0.1530	0.0085	0.7160	0.5096	0.1256	0.0917	0.0630	0.0460	0.0289	0.0276
Product 2	0.2081	-12.6558	0.0324	0.2634	0.3255	0.6412	0.3608	0.0298	0.0241	0.1576	0.0735
Product 3	0.0332	0.0786	-15.2652	0.1789	0.0040	0.3502	6.4654	0.0039	0.0001	0.0089	0.9804
Product4	0.2561	0.0743	0.0194	-5.3782	0.3296	0.2572	0.6071	0.0515	0.0160	0.0249	0.1452
Product 5	0.1969	0.1011	0.0005	0.3491	-3.7918	0.1222	0.0182	0.0329	0.0781	0.0231	0.0030
Product 6	0.1046	0.4534	0.0779	0.6630	0.2504	-11.4914	1.2371	0.0409	0.0132	0.1912	0.2525
Product 7	0.0173	0.0495	0.3801	0.3302	0.0073	0.2295	-7.2039	0.0057	0.0002	0.0065	1.0330
Product 8	0.1716	0.0559	0.0027	0.4106	0.2243	0.1274	0.1011	-5.4275	0.1568	0.0972	0.0762
Product 9	0.0962	0.0376	0.0001	0.0841	0.4119	0.0335	0.0016	0.1109	-4.1643	0.0370	0.0023
Product 10	0.0826	0.3291	0.0068	0.2050	0.1726	0.6596	0.1037	0.1110	0.0493	-9.5316	0.0795
Product 11	0.0226	0.0422	0.2151	0.3090	0.0059	0.1939	4.0385	0.0220	0.0009	0.0226	-10.0816
outside	0.0239	0.0148	0.0020	0.0465	0.0497	0.0159	0.0344	0.0083	0.0137	0.0061	0.0117
variation	15	31	6574	15	127	41	4019	29	1562	31	446

Note: 1. Let i be the index for row and j for column; the entry in cell (i, j) gives the percentage change in the market share of product i with respect to the percentage change in the price of product j. 2. "variation" is the ratio of the maximum value over the minimum value of cross-price elasticities in each column.

Table 2.10: Medians of absolute change in market share w.r.t. percentage change in price in the full model

Before	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	Product 7
Product 1	-45.084	1.059	0.115	7.107	4.920	1.540	1.046
Product 2	0.561	-26.033	0.090	0.870	0.844	2.047	0.838
Product 3	0.051	0.075	-17.062	0.361	0.007	0.490	9.942
Product 4	5.962	1.407	0.698	-101.596	9.341	6.492	18.964
Product 5	5.590	1.788	0.018	12.468	-84.070	3.729	0.516
Product 6	0.804	2.031	0.583	4.008	1.711	-52.921	7.929
Product 7	0.459	0.689	10.089	9.810	0.200	6.620	-102.707

After	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	Product 7	Product 8	Product 9	Product 10	Product 11
Product 1	-35.257	0.872	0.050	3.883	2.748	0.695	0.594	0.345	0.244	0.160	0.147
Product 2	0.438	-27.657	0.067	0.566	0.677	1.418	0.728	0.064	0.057	0.320	0.149
Product 3	0.021	0.056	-11.109	0.129	0.003	0.202	4.663	0.002	0.000	0.006	0.651
Product4	3.654	1.061	0.294	-69.837	4.536	3.547	8.951	0.814	0.242	0.365	2.222
Product 5	3.456	1.686	0.008	6.084	-66.359	1.957	0.292	0.558	1.392	0.382	0.056
Product 6	0.374	1.514	0.261	1.996	0.833	-36.621	4.150	0.131	0.045	0.653	0.783
Product 7	0.266	0.658	5.085	4.242	0.104	3.479	-91.871	0.078	0.002	0.093	14.519
Product8	0.428	0.161	0.007	1.045	0.558	0.302	0.218	-12.636	0.390	0.275	0.185
Product 9	0.375	0.174	0.000	0.397	1.714	0.129	0.006	0.470	-17.551	0.156	0.009
Product 10	0.113	0.456	0.009	0.275	0.226	0.849	0.142	0.154	0.072	-12.344	0.103
Product 11	0.088	0.174	0.910	1.374	0.026	0.926	19.243	0.090	0.004	0.086	-38.699

Note: Let i be the index for row and j for column; the entry in cell (i, j) gives the change in the market share of product i (multiplied by 10^3) with respect to the percentage change in the price of product j.

Table 2.11: Average retail prices and own-price elasticities in the logit and full models

	Before			After		
	price	**logit**	**rcm**	**price**	**logit**	**rcm**
Product 1	11.74 [6]	0.6750 [6]	5.8499 [4]	11.97 [8]	0.7573 [8]	6.2384 [7]
Product 2	22.11 [4]	1.2897 [4]	11.5098 [2]	23.78 [3]	1.5175 [3]	12.6558 [2]
Product 3	26.62 [2]	1.5484 [2]	13.7597 [1]	28.52 [1]	1.8228 [1]	15.2652 [1]
Product 4	13.67 [5]	0.7720 [5]	5.3201 [5]	12.5 [7]	0.7970 [7]	5.3782 [9]
Product 5	10.38 [7]	0.5831 [7]	3.7497 [7]	9.51 [10]	0.6013 [10]	3.7918 [11]
Product 6	22.29 [3]	1.3285 [3]	10.9297 [3]	22.22 [4]	1.4505 [4]	11.4914 [3]
Product 7	26.73 [1]	1.5617 [1]	4.8731 [6]	26.61 [2]	1.7207 [2]	7.2039 [6]
Product 8	—	—	—	9.62 [9]	0.6510 [9]	5.4275 [8]
Product 9	—	—	—	7.74 [11]	0.5001 [11]	4.1643 [10]
Product 10	—	—	—	16.88 [6]	1.1488 [6]	9.5316 [5]
Product 11	—	—	—	20.24 [5]	1.3760 [5]	10.0816 [4]

Note: The numbers in the brackets denote the descending orders of the entries in each column.

Table 2.12: Estimated marginal costs and observed average acquisition costs

	Before				After			
	aac	**mc**	**diff**	**ratio**	**aac**	**mc**	**diff**	**ratio**
product 1	9.29	7.64	-1.65	0.82	9.15	8.01	-1.14	0.88
product 2	17.51	16.83	-0.68	0.96	18.41	18.64	0.23	1.01
product 3	21.00	16.12	-4.88	0.77	21.95	19.15	-2.80	0.87
product 4	10.98	8.41	-2.57	0.77	9.88	7.52	-2.36	0.76
product 5	8.39	6.03	-2.36	0.72	7.43	5.28	-2.15	0.71
product 6	17.62	15.86	-1.76	0.90	16.66	16.03	-0.63	0.96
product 7	21.03	15.68	-5.35	0.75	20.30	16.91	-3.39	0.83
product 8	—	—	—	—	4.12	5.68	1.56	1.38
product 9	—	—	—	—	3.54	4.42	0.88	1.25
product 10	—	—	—	—	7.82	12.01	4.20	1.54
product 11	—	—	—	—	9.38	11.72	2.34	1.25

Note: 1. “diff” is mc minus acc. 2. “ratio” is mc over aac.

Table 2.13: Expected consumer surplus (CS) before and after the entry of SB’s

	Before	**After**	**change**	**% change**
Avg expected CS per unit demand (¢)	1.1	1.21	0.12	10.48
Avg expected CS per market ($)	569.85	625.51	55.66	9.77

Note: 1. “change” is calculated by (After-Before). 2. “% change” is calculated by (change/Before*100).

Table 3.1: Shares of different package sizes

package size	market share	purchase occasion share	feature/display share
3.25 oz	0.22%	0.18%	0.13%
6.5 oz	89.58%	91.40%	99.48%
9.25 oz	4.45%	3.82%	0.39%
12.5 oz	4.20%	2.97%	0.00%
3-pack-3.25 oz	1.55%	1.63%	0.00%
total	100%	100%	100%

Note: 1. "market share" is the quantity (measured by oz) sold of a particular package size over the total quantity sold of the whole product category. 2. "purchase occasion share" is the number of purchase occasions of a particular package size over the total number of purchase occasions. 3. "feature/display share" is the number of weeks of a particular package size on feature/display over the sum of the number of weeks of each package size on feature/display.

Table 3.2: Shares and prices of the dominant brands

brand	market share	purchase occasion share	feature/display share	avg shelf price
1: StarKist	42.17%	45.28%	56.94%	0.72
2: CKN	26.42%	26.19%	26.66%	0.78
3: 3 Diamond	2.18%	2.75%	6.39%	0.75
4: PL	8.37%	7.37%	5.96%	0.67
total	79.15%	81.59%	95.95%	

Note: "avg shelf price" is the average retail price of a particular brand across stores and sample weeks and is measured by dollar.

Table 3.3: Aggregate demand with respect to retail price discount

	# obs	# qty	avg	% obs	% qty
Discount < 10%	6715	431060	64.19	71.38	20.02
10% <= discount < 20%	414	42069	101.62	4.40	1.95
20% <= discount < 30%	832	361891	434.97	8.84	16.81
30% <= discount < 40%	541	398688	736.95	5.75	18.52
40% <= discount < 50%	289	446642	1545.47	3.07	20.75
50% <= discount < 60%	323	294406	911.47	3.43	13.68
Discount >= 60%	293	177908	607.19	3.11	8.26
total	9407	2152664	228.84	100.00	100.00

Note: 1. Each observation is a brand-store-week combination. 2. Quantity is measured by cans. 3. "discount" = $\frac{shelf - regular}{regular}\%$.

Table 3.4: Distributions of households' sale and non-sale purchase quantities

distribution of sale purchase quantity				distribution of non-sale purchase quantity			
	# purchase	%	cum %		# purchase	%	cum %
1 can	2869	28.60	28.60	**1 can**	1341	52.63	52.63
2 cans	5365	53.48	82.09	**2 cans**	978	38.38	91.01
3 cans	1008	10.05	92.13	**3 cans**	133	5.22	96.23
4 cans	571	5.69	97.83	**4 cans**	73	2.86	99.10
5 cans	53	0.53	98.36	**5 cans**	5	0.20	99.29
6 cans	101	1.01	99.36	**6 cans**	11	0.43	99.73
more	64	0.64	100.00	**more**	7	0.27	100.00
total	10031	100.00		**total**	2548	100.00	

	sale	non-sale	total
# purchase	10031	2548	12579
% purchase	79.74	20.26	100

Table 3.5: Distributions of households' purchase frequencies and quantities

# purchases	# hh	%	cum %	# quantities	# purchases	%	cum %
1 purchase	14	0.79	0.79	**1 can**	4210	33.47	33.47
2 purchases	313	17.57	18.36	**2 cans**	6343	50.43	83.89
3 purchases	252	14.15	32.51	**3 cans**	1141	9.07	92.96
4 purchases	193	10.84	43.35	**4 cans**	644	5.12	98.08
5 purchases	162	9.10	52.44	**5 cans**	58	0.46	98.55
6 purchases	137	7.69	60.13	**6 cans**	112	0.89	99.44
7 purchases	99	5.56	65.69	**7 cams**	14	0.11	99.55
8 purchases	100	5.61	71.31	**8 cans**	28	0.22	99.77
9 purchases	81	4.55	75.86	**9 cans**	2	0.02	99.79
10 purchases	67	3.76	79.62	**10 cans**	16	0.13	99.91
11 purchases	54	3.03	82.65	**11 cans**	1	0.01	99.92
12 purchases	41	2.30	84.95	**12 cans**	3	0.02	99.94
13 purchases	42	2.36	87.31	**more**	7	0.06	100.00
14 purchases	25	1.40	88.71	**total**	12579	100.00	
15 purchases	37	2.08	90.79				
16 purchases	28	1.57	92.36				
17 purchases	24	1.35	93.71				
18 purchases	23	1.29	95.00				
19 purchases	18	1.01	96.01				
20 purchases	15	0.84	96.86				
more	56	3.14	100.00				
total	1781	100.00					

Table 3.6: Sale-purchase ratio, purchase duration, and purchase quantity for households with the same purchase frequency

purchase frequency	# hh	% hh	Ratio	Duration	Quantity
2 purchases	313	17.71%	0.81	16.44	2.05
3 purchases	252	14.26%	0.78	14.55	2.03
4 purchases	193	10.92%	0.77	12.44	2.00
5 purchases	162	9.17%	0.81	11.24	2.09
6 purchases	137	7.75%	0.81	10.20	1.92
7 purchases	99	5.60%	0.77	9.40	1.87
8 purchases	100	5.66%	0.79	8.62	1.88
9 purchases	81	4.58%	0.84	8.52	2.02
10 purchases	67	3.79%	0.81	8.36	1.93
11 purchases	54	3.06%	0.80	7.73	1.92
12 purchases	41	2.32%	0.76	7.39	1.86
13 purchases	42	2.38%	0.85	7.27	1.93
14 purchases	25	1.41%	0.84	6.80	2.31
15 purchases	37	2.09%	0.84	6.59	1.98
>= 16 purchases	164	9.28%	0.79	5.23	1.91
total	1767	100.00%	0.80	11.19	1.99

Note: 1. “Ratio” is defined as the number of sale purchases to the number of total purchases for a household. 2. “Duration” is the interval between two purchase occasions, measured by weeks.

Table 3.7: Durations to the next purchase and from the previous purchase for sale and non-sale purchases

household sample	sale next	non-sale next	diff	sale previous	non-sale previous	diff
>= 2 purchases	8.57	7.02	1.55	8.53	7.10	1.43
>= 3 purchases	8.33	6.74	1.59	8.28	6.89	1.39
>= 4 purchases	8.03	6.28	1.75	7.95	6.54	1.41
>= 5 purchases	7.72	6.05	1.67	7.65	6.25	1.40
>= 6 purchases	7.44	5.74	1.70	7.36	6.00	1.35
>= 7 purchases	7.16	5.54	1.62	7.05	5.90	1.15
>= 8 purchases	6.96	5.31	1.64	6.85	5.67	1.19
>= 9 purchases	6.72	5.22	1.50	6.63	5.53	1.10
>= 10 purchases	6.49	5.06	1.43	6.40	5.35	1.06
>= 11 purchases	6.24	4.87	1.37	6.15	5.19	0.95
>= 12 purchases	6.03	4.74	1.29	5.95	5.02	0.93
>= 13 purchases	5.87	4.54	1.33	5.82	4.73	1.09
>= 14 purchases	5.65	4.41	1.24	5.61	4.54	1.07
>= 15 purchases	5.54	4.25	1.29	5.49	4.44	1.04
>= 16 purchases	5.32	4.11	1.21	5.26	4.32	0.93
>= 17 purchases	5.19	4.00	1.18	5.12	4.21	0.90
>= 18 purchases	5.07	3.78	1.30	4.99	4.02	0.97

within sale-nonsale difference	average	# hh
duration to next purchase	1.2200	711
duration from previous purchase	0.8292	676

Note: “diff” is sale next (previous) minus non-sale next (previous).

Table 3.8: Conditional probability that the previous purchase is non-sale

household sample	# sale(t-1) & sale(t) purchase	# nonsale(t-1) & sale(t) purchase	% (NS(t-1) | S(t))	# sale(t-1) & nonsale(t) purchase	# nonsale(t-1) & nonsale(t) purchase	% (NS(t-1) | NS(t))
>= 2 purchases	7495	1157	13.37	1049	1097	51.12
>= 3 purchases	7274	1118	13.32	1022	1071	51.17
>= 4 purchases	6947	1047	13.10	957	1030	51.84
>= 5 purchases	6564	981	13.00	899	958	51.59
>= 6 purchases	6116	905	12.89	833	900	51.93
>= 7 purchases	5625	831	12.87	772	841	52.14
>= 8 purchases	5231	765	12.76	715	764	51.66
>= 9 purchases	4751	690	12.68	648	686	51.42
>= 10 purchases	4272	622	12.71	583	650	52.72
>= 11 purchases	3849	553	12.56	519	603	53.74
>= 12 purchases	3471	495	12.48	470	548	53.83
>= 13 purchases	3177	442	12.21	420	494	54.05
>= 14 purchases	2794	397	12.44	381	457	54.53
>= 15 purchases	2544	375	12.85	361	424	54.01
>= 16 purchases	2156	328	13.20	319	383	54.56
>= 17 purchases	1830	294	13.84	288	354	55.14
>= 18 purchases	1534	260	14.49	255	333	56.63

within sale-nonsale difference	avg	# hh
cond. prob. of a previous non-sale purchase	8.45%	676

Table 3.9: Duration effects during the sale and non-sale periods

log(qty)	Coef.	Std.	t test	log(qty)	Coef.	Std.	t test
log(price)	-3.158	0.043	-73.730	**log(price)**	-2.832	0.048	-59.250
feature	0.045	0.032	1.430	**feature**	0.043	0.031	1.380
display	0.025	0.044	0.560	**display**	0.016	0.044	0.360
lent	0.004	0.024	0.170	**lent**	0.012	0.024	0.500
constant	2.787	0.057	49.010	**constant**	2.916	0.057	51.260
duration	0.004	0.003	1.400	**duration-sale**	0.139	0.010	14.530
				duration-nonsale	-0.003	0.003	-1.200
# of obs	9225			**# of obs**	9225		
Adj R^2	0.714			**Adj R^2**	0.739		

Note: 1.Each observation is a brand-week-store combination. 2. The covariate "duration" is the number of weeks since last sale of any brand in the same store. 3. The covariate "duration-sale" is zero when the brand is not on sale and equals "duration" when it is on sale that week; similarly, the covariate "duration-nonsale" is zero when the brand is on sale and equals "duration" when it is not on sale that week. 4. Store- and brand-specific dummy variables are included as covariates in the regressions as well; the results are not reported here to save space.

Table 3.10: Frequencies of at least one brand on sale

	sale
Chain 1	42.28%
Chain 2	54.59%
Chain 3	47.80%
Chain 4	98.17%

Note: A sale is defined when $\frac{shelf - regular}{regular} \geq 20\%$

Table 3.11: Durations to the next purchase and from the previous purchase for sale and non-sale purchases on a chain basis

	duration to next purchase			duration from previous pruchase		
	sale	non-sale	diff	sale	non-sale	diff
chain 1	7.85	6.05	1.80	7.57	6.44	1.14
chain 2	7.68	6.29	1.39	7.96	5.47	2.49
chain 3	7.35	6.37	0.98	7.40	6.04	1.36
chain 4	7.21	6.11	1.10	7.08	7.44	-0.36

Note: “diff” is sale minus non-sale.

Table 3.12: Conditional probability that the previous purchase is non-sale on a chain basis

	# s(t)	# (ns(t-1) | s(t))	% (ns(t-1) | s(t))	# ns(t)	# (ns(t-1) | ns(t))	% (ns(t-1) | ns(t))
chain 1	1001	258	25.8%	648	404	62.3%
chain 2	1178	229	19.4%	495	315	63.6%
chain 3	1428	153	10.7%	240	99	41.3%
chain 4	1949	111	5.7%	198	83	41.9%

Table 3.13: Duration effects during the sale and non-sale periods on a chain basis

chain 1			
log(qty)	**Coef.**	**Std.**	**t test**
log(price)	-3.437	0.084	-40.99
duration-sale	0.047	0.009	5.33
duration-nonsale	0.001	0.002	0.34
# of obs	3956		
Adj R^2	0.790		
chain 2			
log(qty)	**Coef.**	**Std.**	**t test**
log(price)	-3.577	0.128	-27.87
duration-sale	0.251	0.026	9.58
duration-nonsale	0.083	0.012	7.1
# of obs	1711		
Adj R^2	0.785		
chain 3			
log(qty)	**Coef.**	**Std.**	**t test**
log(price)	-1.848	0.085	-21.63
duration-sale	0.508	0.042	12.21
duration-nonsale	-0.052	0.025	-2.06
# of obs	2365		
Adj R^2	0.685		
chain 4			
log(qty)	**Coef.**	**Std.**	**t test**
log(price)	-5.420	0.273	-19.85
duration-sale	0.300	0.176	1.71
duration-nonsale	0.095	0.156	0.61
# of obs	1193		
Adj R^2	0.741		

Table 3.14: Estimation results for the proportional hazard model

household sample	sale(t)	feature(t)	display(t)	sale(t-1)	hhsize	log(hhinc)
>= 2 purchases	-0.133	0.021	0.057	-0.172	0.068	0.033
>= 3 purchases	-0.130	0.013	0.054	-0.178	0.066	0.030
>= 4 purchases	-0.126	0.015	0.049	-0.207	0.061	0.028
>= 5 purchases	-0.137	0.019	0.060	-0.203	0.063	0.023
>= 6 purchases	-0.134	0.013	0.056	-0.223	0.070	0.005
>= 7 purchases	-0.113	0.006	0.087	-0.236	0.069	-0.014
>= 8 purchases	-0.121	0.000	0.060	-0.244	0.060	-0.003
>= 9 purchases	-0.124	0.007	0.065	-0.237	0.053	-0.003
>= 10 purchases	-0.126	0.018	0.074	-0.240	0.054	0.000
>= 11 purchases	-0.121	0.029	0.058	-0.241	0.046	-0.022
>= 12 purchases	-0.114	0.018	0.043	-0.240	0.049	-0.034
>= 13 purchases	-0.136	0.011	0.021	-0.250	0.040	-0.032
>= 14 purchases	-0.144	0.007	-0.010	-0.240	0.037	-0.045
>= 15 purchases	-0.136	0.013	-0.011	-0.266	0.034	-0.048
>= 16 purchases	-0.112	-0.003	-0.021	-0.264	0.030	-0.039

t test	sale(t)	feature(t)	display(t)	sale(t-1)	hhsize	log(hhinc)
>= 2 purchases	-5.443	1.084	2.180	-7.649	9.804	2.327
>= 3 purchases	-5.209	0.650	2.051	-7.793	9.435	2.048
>= 4 purchases	-4.829	0.734	1.815	-8.660	8.449	1.839
>= 5 purchases	-5.053	0.896	2.124	-8.266	8.589	1.490
>= 6 purchases	-4.797	0.592	1.948	-8.800	8.580	0.329
>= 7 purchases	-3.927	0.260	2.820	-8.914	8.157	-0.825
>= 8 purchases	-4.044	0.010	1.879	-8.690	6.761	-0.195
>= 9 purchases	-3.993	0.289	1.919	-8.067	5.592	-0.169
>= 10 purchases	-3.839	0.689	2.072	-7.715	5.237	0.022
>= 11 purchases	-3.448	1.036	1.533	-7.343	4.199	-1.003
>= 12 purchases	-3.097	0.595	1.079	-6.845	4.220	-1.412
>= 13 purchases	-3.474	0.364	0.511	-6.782	3.198	-1.245
>= 14 purchases	-3.466	0.224	-0.234	-6.214	2.826	-1.704
>= 15 purchases	-3.174	0.381	-0.240	-6.681	2.476	-1.828
>= 16 purchases	-2.468	-0.075	-0.434	-6.303	2.121	-1.376

Table 3.15: Purchase ratios and quantities of sale/non-sale purchases

household sample	# household	# sale purchase	# non-sale purchase	% sale purchase	% non-sale purchase	sale qty	non-sale qty	diff
>= 2 purchases	1767	8544	2254	79.13	20.87	2.02	1.63	0.40
>= 3 purchases	1454	8296	2189	79.12	20.88	2.02	1.63	0.40
>= 4 purchases	1202	7904	2077	79.19	20.81	2.02	1.62	0.39
>= 5 purchases	1009	7463	1939	79.38	20.62	2.01	1.62	0.40
>= 6 purchases	847	6949	1805	79.38	20.62	2.00	1.62	0.38
>= 7 purchases	710	6397	1672	79.28	20.72	2.00	1.62	0.38
>= 8 purchases	611	5946	1529	79.55	20.45	2.00	1.63	0.38
>= 9 purchases	511	5399	1376	79.69	20.31	2.01	1.64	0.37
>= 10 purchases	430	4855	1272	79.24	20.76	2.00	1.65	0.34
>= 11 purchases	363	4368	1156	79.07	20.93	1.99	1.65	0.34
>= 12 purchases	309	3941	1043	79.07	20.93	2.00	1.64	0.36
>= 13 purchases	268	3597	936	79.35	20.65	2.01	1.65	0.36
>= 14 purchases	226	3175	854	78.80	21.20	2.01	1.65	0.37
>= 15 purchases	201	2905	799	78.43	21.57	1.99	1.56	0.43
>= 16 purchases	164	2475	711	77.68	22.32	1.99	1.54	0.44
>= 17 purchases	136	2118	648	76.57	23.43	1.95	1.54	0.41
>= 18 purchases	112	1789	593	75.10	24.90	1.95	1.55	0.39

Appendix C

Figures

Figure 1.1: Retail price patterns of StarKist and CKN in a Chain #1 store

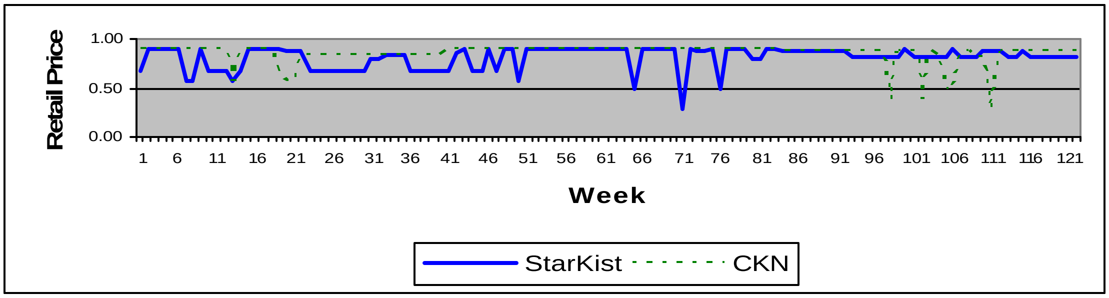


Figure 1.2: Retail price patterns of StarKist and CKN in a Chain #2 store

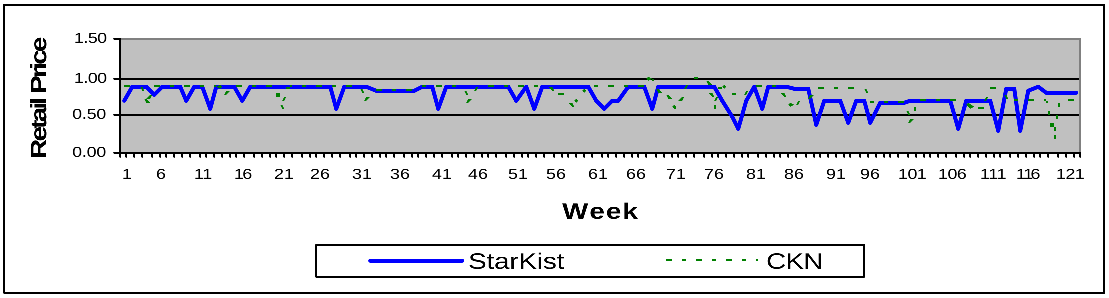


Figure 1.3: Retail price patterns of StarKist and CKN in a Chain #3 store

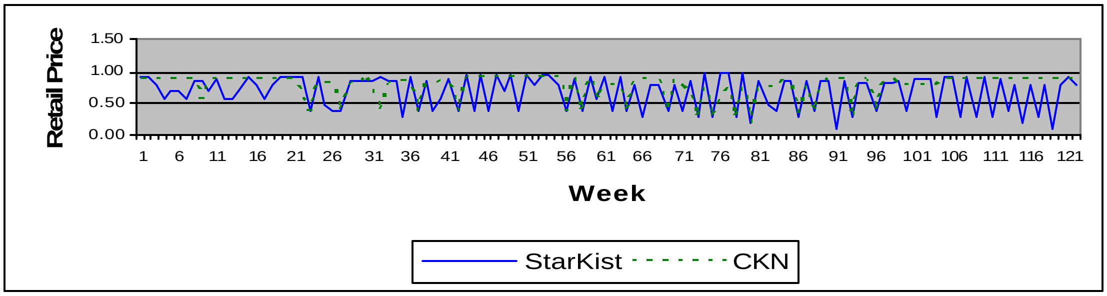


Figure 1.4: Retail price patterns of StarKist and CKN in a Chain #4 store

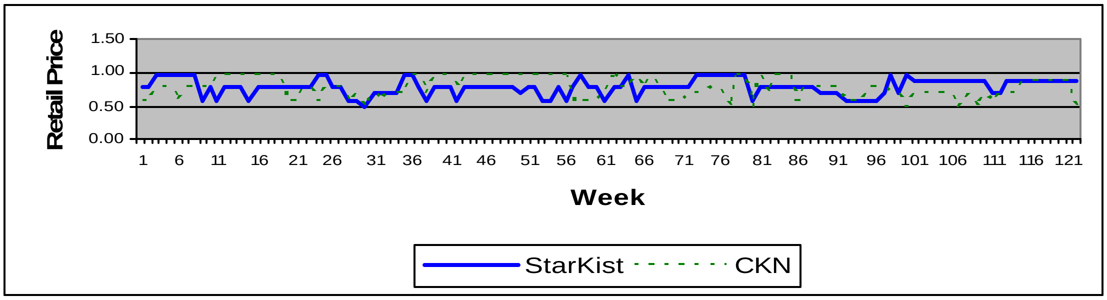


Figure 1.5: Retail price patterns of StarKist in four supermarket chains

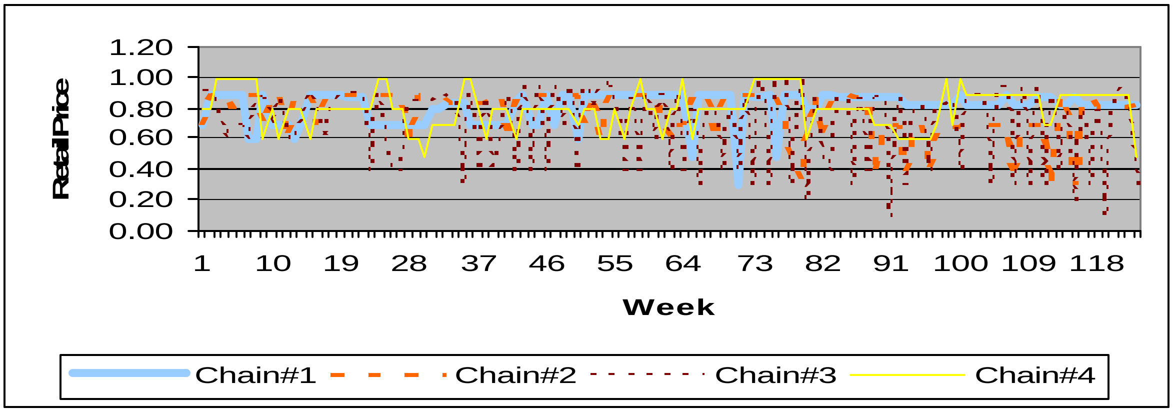


Figure 1.6: Retail price patterns of CKN in four supermarket chains

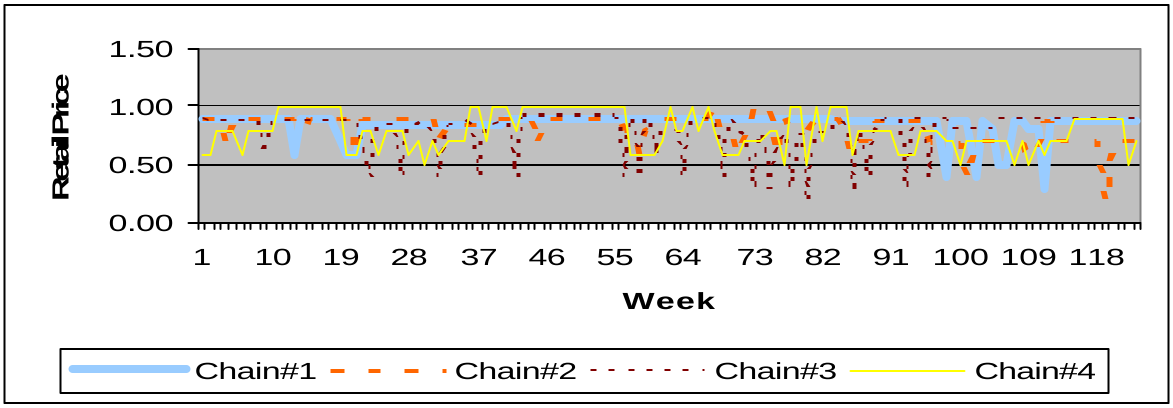


Figure 1.7: Utility from the outside good w.r.t. the elapsed time since last purchase

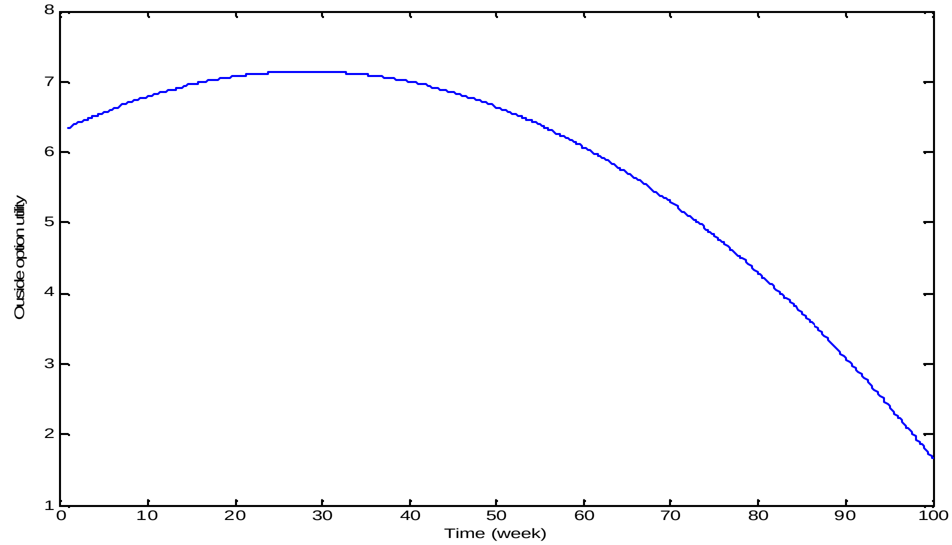


Figure 1.8: Logarithm of simulated category demand

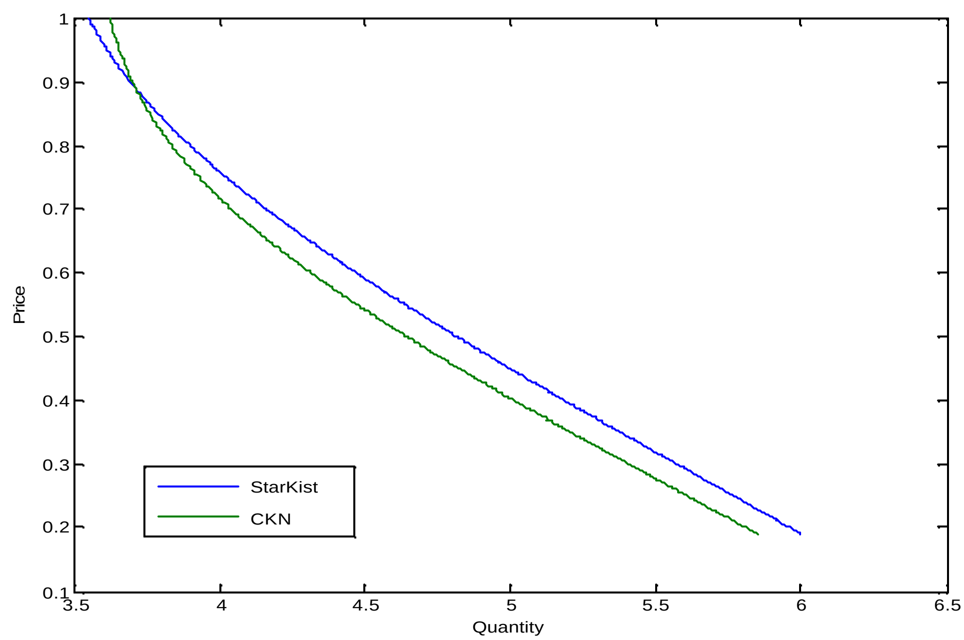
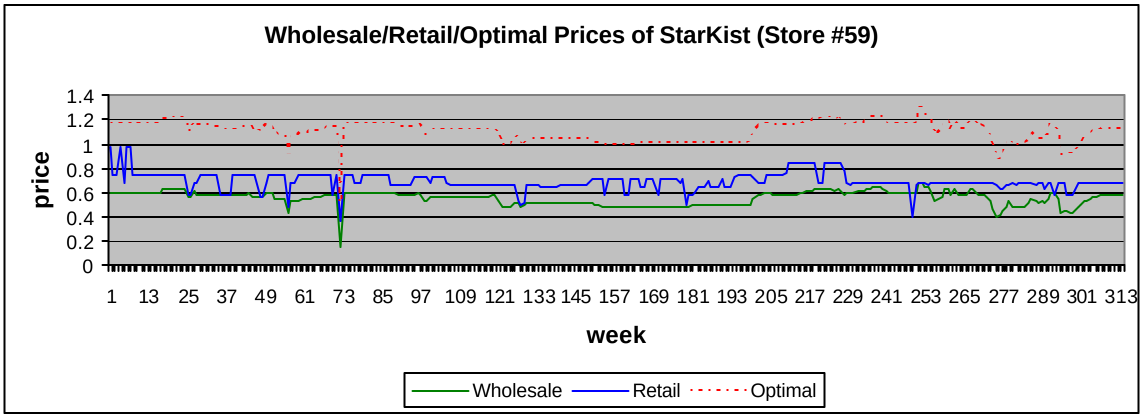


Figure 1.9: Price patterns of StarKist and CKN in Dominick's store #59



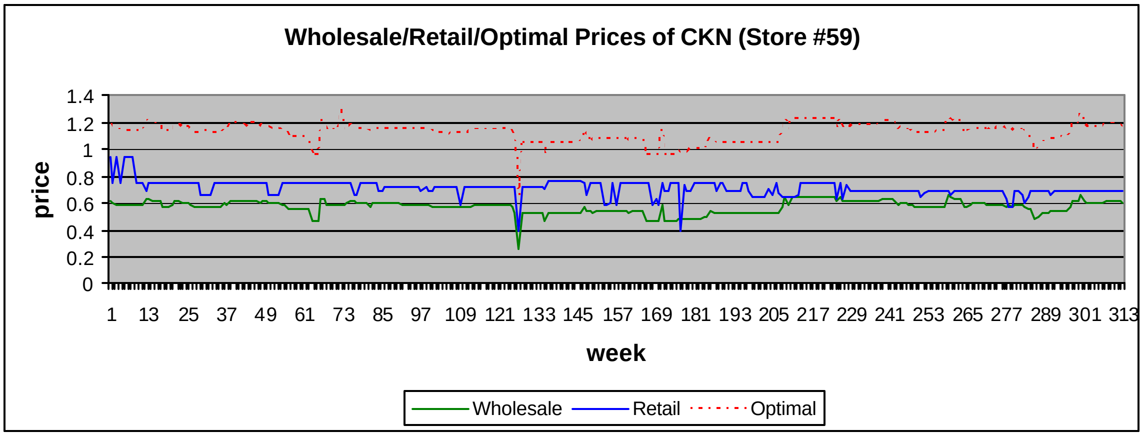
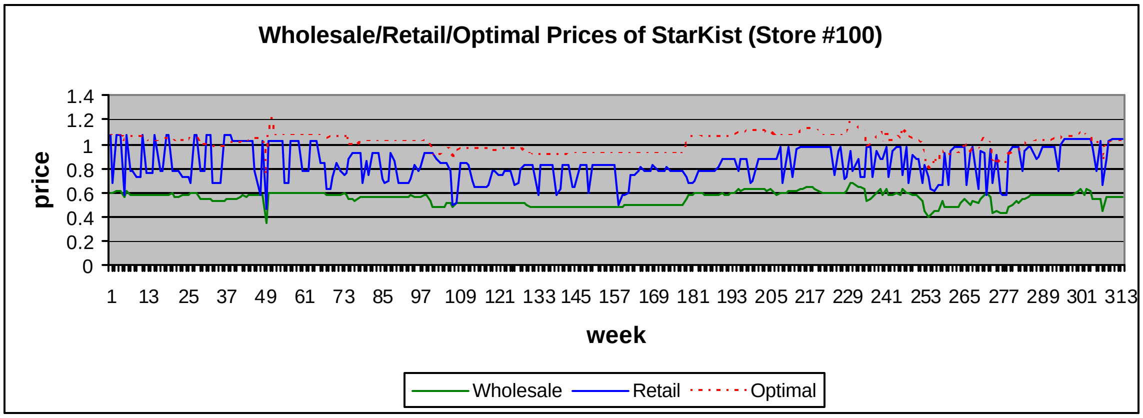


Figure 1.10: Price patterns of StarKist and CKN in Dominick's store #100



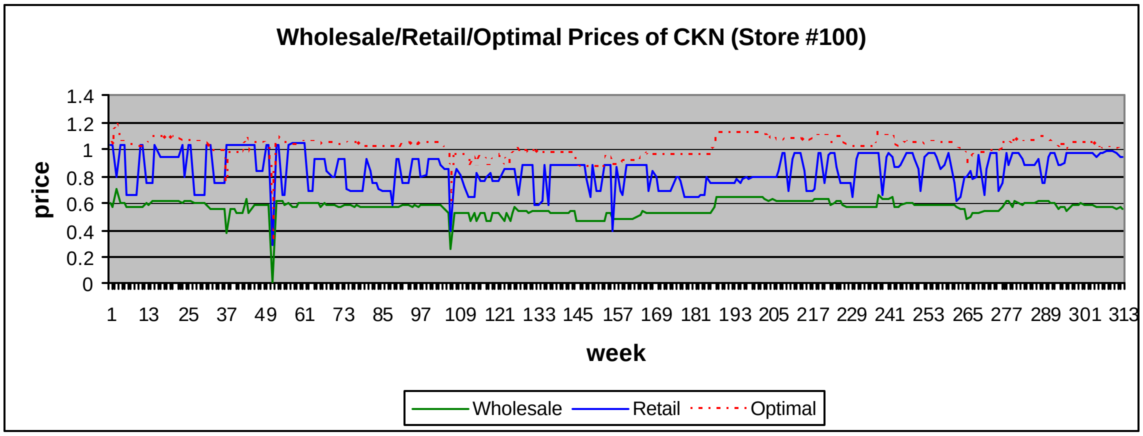


Figure 3.1: An example of retail price patterns of StarKist and CKN

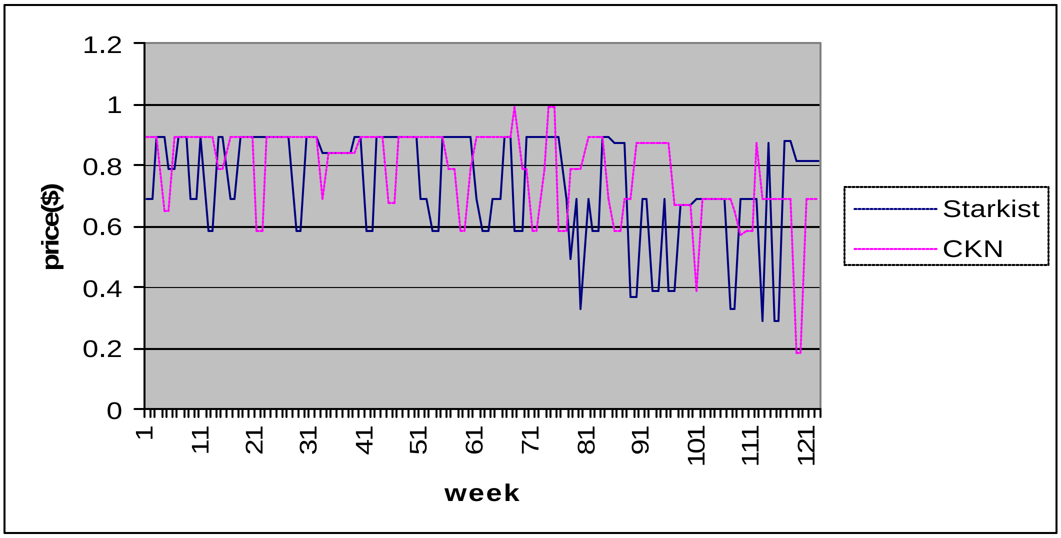
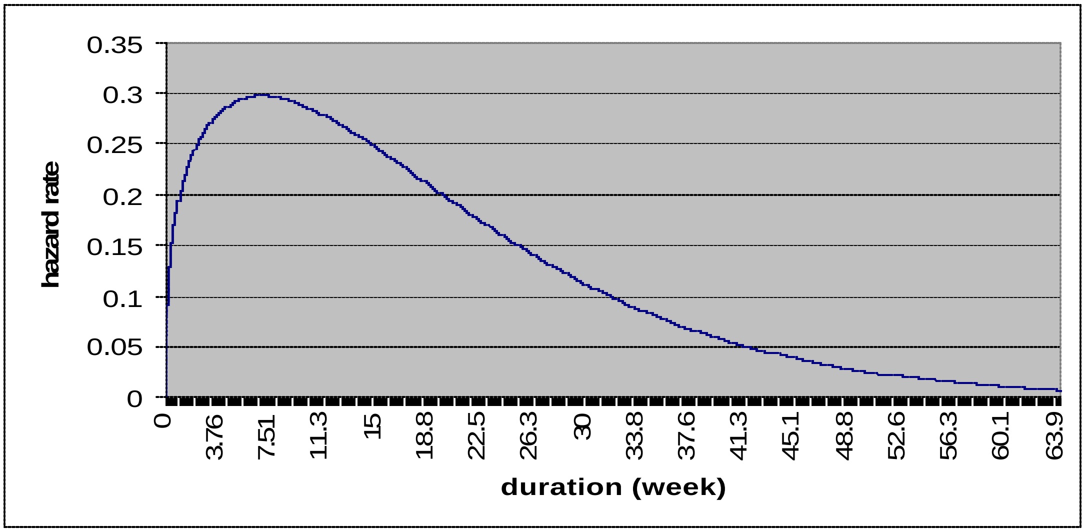


Figure 3.2: Baseline expo-power hazard function



Bibliography

[1] Aguirregabiria, Victor (1999), "The dynamics of markups and inventories in retailing firms," Review of Economic Studies, 66, 275-308.

[2] Ailawadi, Kusum (2001), "The retail power-performance conundrum: What have we learned," Journal of Retailing, 77, 299-318.

[3] Ailawadi, Kusum and Bari Harlam (2002), "The effect of store brands on retail margins: An empirical analysis," Working paper.

[4] Allenby, Greg and Peter Lenk (1994), "Modeling household purchase behavior with logistic normal regression," Journal of the American Statistical Association, 89(428), 1218-1231.

[5] Allenby, Greg and Peter Rossi (1999), "Marketing Models of Consumer Heterogeneity," Journal of Econometrics, 89, 57-78.

[6] Allenby, Greg, Thomas Shively, Sha Yang, and Mark Garratt (2004), "A choice model for packaged goods: dealing with discrete quantities and quantity discounts," Marketing Science, 23(1), 95-108.

[7] Amir, Rabah (2005), "Supermodularity and complementarity in Economics: A elementary survey," Southern Economic Journal, 71(93), 636-660.

[8] Andrews, Donald (1999), "Consistent moment selection procedures for generalized method of moments estimation," Econometrica, 67(3), 543-564.

[9] Anton, James and Gopal Das Varma (2005), "Storability, market structure, and demand-shift incentives," RAND Journal of Economics, 36(3), 520-543.

[10] Asker, John (2004), "Measuring advantages from exclusive dealing," Working paper.

[11] Barsky, Robert, Shantanu Dutta, and Daniel Levy (2000), "What can the price gap between branded and generic tell us about markups," Working paper.

[12] Bell, David, Jeongwen Chiang, and V. Padmanabhan (1999), "The decomposition of promotional response: An empirical generalization," Marketing Science, 18(4), 504-526.

[13] Berges-Sennou, Fabian, Philippe Bontems, and Vincent Requillart (2004), "Economic impact of the development of private labels," Working paper.

[14] Berry, Steven (1994), "Estimating discrete-choice models of product differentiation," Rand Journal of Economics, 25, 242-262.

[15] Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium," Econometrica, 63(4), 841-890.

[16] Berry, Steven, James Levinsohn, and Ariel Pakes (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," Journal of Political Economy, 112(1), 68-105.

[17] Blattberg, Robert, Gary Eppen, and Joshua Lieberman (1981), "A theoretical and empirical evaluation of price deals for consumer nondurables," Journal of Marketing, 45(Winter), 116-129.

[18] Blattberg, Robert and Alan Levin (1987), "Modeling the effectiveness and profitability of trade promotions," Marketing Science, 6(2), 124-146.

[19] Blattberg, Robert, Richard Briesch, and Edward Fox (1995), "How promotions work," Marketing Science, 14(3), G122-G132.

[20] Bonfrer, Andre and Pradeep Chintagunta (2004), "Store brands: who buys them and what happens to retail prices when they are introduced?," Review of Industrial Organization, 24, 195-218.

[21] Borin, Norm and Paul Farris (1990), "An empirical comparison of direct product profitability and existing measures of SKU productivity," Journal of Retailing, 66(3), 297-314.

[22] Bucklin, Randolph, Sunil Gupta, and S. Siddarth (1998), "Determining segmentation in sales response across consumer purchase behaviors," Journal of Marketing Research, 35(2), 189-197.

[23] Chevalier, Judith, Anil Kashyap, and Peter Rossi (2003), "Why don't prices rise during periods of peak demand? - Evidence from scanner data," The American Economic Review, 93(1), 15-37.

[24] Chiang, Jeongwen (1991), "A simultaneous approach to the whether, what and how much to buy questions," Marketing Science, 10(4), 297-315.

[25] Chintagunta, Pradeep (1993), "Investigating purchase incidence, brand choice and purchase quantity decisions of households," Marketing Science, 12(2), 184-208.

[26] Chintagunta, Pradeep (2002), "Investigating category pricing behavior at a Retail Chain," Journal of Marketing Research, 39, 141-154.

[27] Chintagunta, Pradeep, Andre Bonfrer, and Inseong Song (2002), "Investigating the effects of store-brand introduction on retailer demand and pricing behavior," Management Science, 48, 1242-1267.

[28] Chintagunta, Pradeep, Jean-Pierre Dube, and Vishal Singh (2003), "Balancing Profitability and Customer Welfare in a Supermarket Chain," Quantitative Marketing and Economics, 1(1), 111-147.

[29] Cohen, Andrew (2004), "Identifying price discrimination when product menus are endogenous," Working paper.

[30] Corstjens, Marcel, and Rajiv Lal (2000), "Building Store Loyalty Through Store Brands," Journal of Marketing Research, 37(3), 281-291.

[31] Dhar, Sanjay and Stephen Hoch (1997), "Why store brand penetration varies by retailer," Marketing Science, 16(3), 208-227.

[32] Draganska, Michaela and Dipak Jain (2005), "Product-line length as a competitive tool," Journal of Economics & Management Strategy, 14(1), 1-28.

[33] Dréze, Xavier and David Bell (2003), "Creating Win-Win Trade Promotions: Theory and Empirical Analysis of Scan-back Trade Deals," Marketing Science, 22(1), 16-39.

[34] Dube, Jean-Pierre (2005), "Product Differentiation and Mergers in the Carbonated Soft Drink Industry," Journal of Economics and Management Strategy, 14(4), 879-904.

[35] Gupta, Sunil (1988), "Impact of sales promotions on when, what, and how much to buy," Journal of Marketing Research, 25(4), 342-355.

[36] Greene, William, *Econometric Analysis*, New Jersey: Prentice-Hall, Inc.

[37] Hanemann, William (1984), "Discrete/continuous models of consumer demand," Econometrica, 52(3), 541-561.

[38] Hansen, Karsten, Vishal Singh, and Pradeep Chintagunta (2006), "Understanding store brand purchase behavior across categories," Marketing Science, 25(1), 75-90.

[39] Hartmann, Wesley (2006), "Intertemporal effects of consumption and their implication for demand elasticity estimates," Quantitative Marketing and Economics, 4(4), 325-349.

[40] Heedre, Harald, Sachin Gupta, and Dick Wittink (2003), "Is 75% of the sales promotion bump due to brand switching? No, only 33% is," Journal of Marketing Research, 40(4), 481-491.

[41] Hendel, Igal (1999), "Estimating multiple-discrete choice models: An application to computerization returns," Review of Economic Studies, 66, 423-446.

[42] Hendel, Igal and Aviv Nevo (2002), "Sales and consumer inventory," forthcoming in The RAND Journal of Economics.

[43] Hendel, Igal and Aviv Nevo (2005), "Measuring the implications of sales and consumer inventory behavior," forthcoming in Econometrica.

[44] Hong, Pilky, Preston McAfee, and Ashish Nayyar (2002), "Equilibrium Price Dispersion with consumer inventories" Journal of Economic Theory, 105, 503-517.

[45] Jain, Dipak and Naufel Vilcassim (1991), "Investigating household purchase timing decisions: a conditional hazard function approach," Marketing Science, 10(1) 1-23.

[46] Keane, Michael (1997), "Modeling heterogeneity and state dependence in consumer choice behavior," Journal of Business & Economic Statistics, 15(3), 310-327.

[47] Kim, Jaehwan, Greg Allenby, Peter Rossi (2002), "Modeling consumer demand for variety," Marketing Science, 23(3), 229-250.

[48] Lal, Rajiv (1990), "Manufacturer trade deals and retail price promotions," Journal of Marketing Research, 27(4), 428-444.

[49] Lal, Rajiv and Miguel Villas-Boas (1998), "Price promotions and trade deals with multiproduct retailers," Management Science, 44(7), 935-949.

[50] McAfee, Preston (1995), "Multiproduct equilibrium price dispersion," Journal of Economic Theory, 67(1), 83-105.

[51] Meza, Sergio and K. Sudhir (2003), "The role of strategic pricing by retailers in the success of store brands," Working paper.

[52] Morton, Fiona and Florian Zettelmeyer (2004), "The strategic positioning of store brands in retailer-manufacturer bargaining," Review of Industrial Organization, 24, 161-194.

[53] Narasimhan, Chakravarthi and Ronald Wilcox (1998), "Private labels and the channel relationship: a cross-category analysis," The Journal of Business, 71(4), 573-600.

[54] Nevo, Aviv (2000), "A practitioner's guide to estimation of random-coefficients Logit models of demand," Journal of Economics & Management Strategy, 9, 513-548.

[55] Nevo, Aviv (2001), "Measuring market power in the ready-to-eat cereal industry," Econometrica, 69, 307-342.

[56] Nevo, Aviv and Konstantinos Hatzitaskos (2005), "Why does the average price of tuna fall during Lent?," Working paper.

[57] Osborne, Matthew (2005), "Consumer learning, habit formation, and heterogeneity: A structural examination," Working paper.

[58] Pauwels, Koen and Shuba Srinivasan (2004), "Who benefits from store brand entry?," Marketing Science, 23(3), 364-390.

[59] Pesendorfer, Martin (2002), "Retail sales: a study of pricing behavior in supermarkets," Journal of Business, 75(1), 33-66.

[60] Petrin, Amil (2002), "Quantifying the Benefits of New Products: The Case of the Minivan," Journal of Political Economics, 110, 705-729.

[61] Putsis, William (1997), "An empirical study of the effect of brand proliferation on private label-national brand pricing behavior," Review of Industrial Organization, 12, 355-371.

[62] Quelch, John and David Harding (1996), "Brands versus provate labels: Fighting to win," Harvard Business Review 74(1), 99-109.

[63] Rossi, Peter and Greg Allenby (1993), "A Bayesian approach to estimating household parameters," Journal of Marketing Research, 30(2), 171-182.

[64] Rossi, Peter, Robert McCulloch, and Greg Allenby (1996), "The value of purchase history data in target marketing," Marketing Science, 15(4), 321-340.

[65] Rossi, Peter and Greg Allenby (2003), "Bayesian statistics and marketing," Marketing Science, 22(3), 304-328.

[66] Sayman, Serdar and Jagmohan Raju (2004), "Investigating the cross-category effects of store brands," Review of Industrial Organization, 24, 129-141.

[67] Silva-Risso, Jorge, Randolph Bucklin, and Donald Morrison (1999), "A decision support system for planning manufacturers' sales promotion calendars," Marketing Science, 18(3), 274-300.

[68] Sobel Joel (1984), "The timing of sales," Review of Economic Studies, 51, 353-368.

[69] Steiner, Robert (2004), "The nature and benefits of national brnad/private label competition," Review of Industrial Organization, 24, 105-127.

[70] Struse, Rudolph (1987), "Approaches to promotion evaluation: a practitioner's viewpoint," Marketing Science, 6(2), 150-151.

[71] Sudhir, K. and Debabrata Talukdar (2004), "Does Store Brand Patronage Improve Store Patronage?," Review of Industrial Organization, 24, 143-160

[72] Tellis, Gerard and Fred Zufryden (1995), "Tackling the retailer decision maze: Which brands to discount, how much, when and why?," Marketing Science, 14(3), 271-299.

[73] Train, Kenneth, *Discrete choice methods with simulation*, Cambridge University Press.

[74] Tyagi, Rafeev (1999), "A characterization of retailer response to manufacturer trade deals," Journal of Marketing Research, 36(4), 510-516.

[75] Varian, Hal (1980), "A model of sales," American Economic Review, 70, 651-388.

[76] Vilcassim, Naufel and Pradeep Chintagunta (1995), "Investigating retailer product category pricing form household scanner panel data," Journal of Retailing, 71(2) 103-128.

[77] Villas-Boas, Sofia and Miguel Villas-Boas (2006), "Learning, forgetting, and sales," CUDARE working paper.

[78] Villas-Boas, Sofia (2006), "Vertical relationships between manufacturers and retailers: inference with limited data," CUDARE working paper.

[79] Walters, Rockney (1989), "An empirical investigation into retailer response to manufacturer trade promotions," Journal of Retailing, 65(2), 253-272.

[80] Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data*, London MIT Press.

Vita

An-Shih Liu was born in Taipei, Taiwan on November 2 1977, the youngest daughter of Yu-Chi Liu and Mei-Yu Lin. She received the Bachelor of Science degree in Money and Banking from National Cheng-Chi Univeristy in June 2000. In August 2002 she entered the graduate program of Economics at The University of Texas at Austin.

Permanent address: 1 University Station C3100
Austin, Texas 78712

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.